DOCUMENT RESUME

ED 107 918 CE 004 036

TITLE Testing Guidelines for Adult Basic Education and High School Equivalency Programs.
INSTITUTION New Jersey State Dept. of Education, Trenton. Office of Adult Basic Education.
PUB DATE 73
NOTE 187p.

EDRS PRICE MF-$0.76 HC-$9.51 PLUS POSTAGE
DESCRIPTORS Achievement Tests; *Adult Basic Education; Adult Education Programs; Adult Students; Diagnostic Tests; Educational Improvement; *Educational Testing; English (Second Language); Guides; *High School Equivalency Programs; Program Improvement; Screening Tests; *Standardized Tests; Student Testing; Test Interpretation; Test Reliability; Test Reviews; *Test Selection; Test Validity

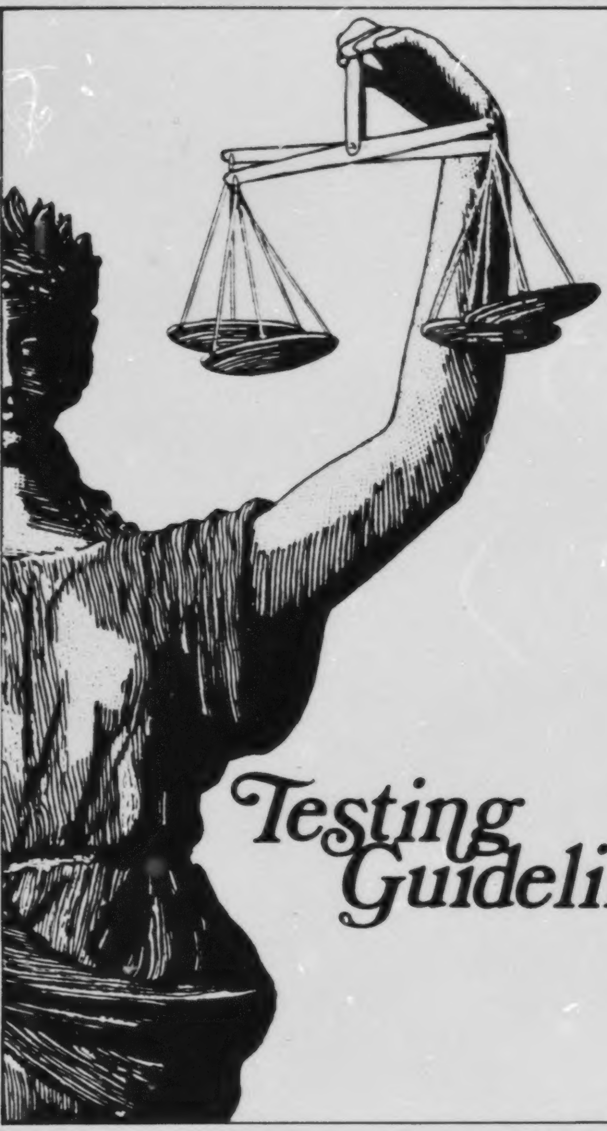
ABSTRACT

The document is a handbook, designed to assist directors, teachers, and counselors in adult basic education (ABE) and high school equivalency programs to plan an overall testing program by offering how-to-do-it ideas in layman terms. It was developed by a group of experienced ABE practitioners to improve testing efforts in adult education programs in New Jersey. Part One, "Testing," offers three areas of basic information: (1) improving educational programs through the use of standardized tests; (2) designing an effective testing program; and (3) testing English as a Second Language (ESL) students. Part Two, "Test Reviews," provides a condensation of pertinent information and suggests specific tests for identifiable purposes. Twenty specific tests of four types are reviewed: sceeening, achievement, diagnostic, and ESL. The test review format is intended to provide a structure for examining and evaluating tests which may be published in the future. The narrative section of the book was also constructed to be of value in the future, with the expectation that testing may be constantly changing as better instruments are developed. (Author/AJ)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED107918

MAR 26 1975



Testing Guidelines

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

for ADULT BASIC EDUCATION and HIGH SCHOOL EQUIVALENCY PROGRAMS

ED107918

NEW JERSEY STATE DEPARTMENT OF EDUCATION

COMMISSIONER OF EDUCATION

DIVISION OF FIELD SERVICES
JOHN ROSSER, DIRECTOR
CATHERINE HAVRILESKY, DEPUTY DIRECTOR

ADULT & CONTINUING EDUCATION
ROCCO GAMBACORTA, DIRECTOR

ADULT BASIC EDUCATION
GEORGE A. SNOW, DIRECTOR

Mr. Calvin J. Hurd, President	Elizabeth
Mrs. Hugh Auchincloss, Vice President	Ridgewood
Mr. Paul J. Christiansen	West Orange
Mrs. Marion G. Epstein	Princeton
Mr. John J. Grossi, Jr.	Westfield
Mrs. Ruth Mancuso	Glassboro
Mrs. M. Patricia O'Hara	Butler
Dr. Allen G. Potter, Jr.	Neshanic Station
Mr. Harry M. Seals	Annandale
Mr. Jack Slater	Paterson
Mr. George F. Smith	Edison
Mrs. Helen A. Zehner	Woodstown

TESTING GUIDELINES

FOR ADULT BASIC EDUCATION
AND HIGH SCHOOL EQUIVALENCY PROGRAMS

Joan Fischer
Director
Adult Education Resource Center
Montclair State College
Upper Montclair, New Jersey

Jane F. Flaherty
Assistant Director
Adult Education Resource Center
Newark State College
Union, New Jersey

Robert H. Arents, Director
Educational Services
WIN Program
Committee Chairman

State of New Jersey
Department of Education
Division of Field Services
Bureau of Adult and Continuing Education
The Office of Adult Basic Education
Trenton, New Jersey

1973 Edition

FOREWORD

This handbook is designed to assist directors, teachers and counselors in adult basic education and high school equivalency programs to plan an overall testing program by offering how-to-do-it ideas in layman terms, and suggesting specific tests for identifiable purposes. It represents many months of coordinated activity directed toward the development of a practical document to improve testing efforts in adult education programs in New Jersey.

Many individuals from the New Jersey State Department of Education and respective staff members from the Adult Education Resource Centers at Glassboro, Jersey City, Newark and Montclair State Colleges are to be complimented for their valuable contributions. Robert H. Arents, Director, Education Services, WIN Program, Chairman of the Committee for Development of Guidelines for Testing in Adult Education, is to be credited with the overall design and development of the booklet. Special recognition must be given to Joan Fischer, Director of the Montclair Adult Education Resource Center, for her persistence in gathering and reviewing multitudinous materials on testing and her diligence in writing the numerous refinement drafts for this booklet. Jane Flaherty, Assistant Director of the Newark Adult Education Resource Center, must also be recognized for her outstanding assistance in gathering, writing and reviewing materials contained in this publication. Without their help and diligence, this booklet would not have been possible.

Kathryn Taschler and Joyce Wood, Assistant Directors at the Jersey City and Glassboro Adult Education Resource Centers, respectively, are to be acknowledged for their suggestions in the area of English as a

Second Language. Finally, Bruno Ciccariello, Assistant Director, Office of Adult Basic Education, must be acknowledged for undertaking the tremendous task of editing the original draft of the booklet.

For their efficiency and conscientiousness in typing the many revisions and test reviews, several secretaries must be commended: Fay Deroian, Betty Cohn and Catherine Millichap of the Montclair Adult Education Resource Center. Finally, the Adult Education Resource Center of Montclair State College is to be thanked for assuming the task of duplicating and distributing the original 1971 edition to ABE and High School Equivalency directors of New Jersey.

Although the testing picture for adults will no doubt be constantly changing as better instruments are being developed and produced by publishers, the narrative section of this handbook is constructed to be valuable by itself for a long time to come. Additionally, the test review format, as well as several of the reviews, will provide a structure for examining future tests and identifying their strengths and weaknesses. Everyone involved in testing or individually prescribing an education program for adult participants in basic education is, therefore, encouraged to utilize it as a guide. Those new to adult education will find Part I of particular assistance for basic information whereas experienced persons in the field will find the test reviews a valuable condensation of pertinent information. The few minutes spent reviewing this guide will be well worthwhile, since the recommendations come from a group of experienced practitioners in the field of Adult Basic Education.

George A. Snow
Director
Office of Adult Basic Education

TABLE OF CONTENTS

PAGE

FOREWORD

PART I: TESTING

SECTION 1: IMPROVING EDUCATIONAL PROGRAMS THROUGH THE USE OF STANDARDIZED TESTS. 1

Types of Useful Tests 2

Using Test Results for Decision-Making. 4

SECTION 2: DESIGNING AN EFFECTIVE TESTING PROGRAM

Technical Considerations. 9

Scheduling Tests and Retests. 13

Preparing Students for Standardized Testing 15

Interpreting Test Scores. 16

Reporting Test Results to the Individual. 18

Criteria of a Testing Program 20

Cautions About the Use of Tests 22

SECTION 3: TESTING ENGLISH AS A SECOND LANGUAGE STUDENTS 23

PART II: TEST REVIEWS

SECTION 4: SCREENING TESTS

Dolch Basic Sight Word Test (DBS) 29

Harris Graded Word List (HGWL). 34

SRA Reading and Arithmetic Indexes (RAI). 38

Slosson Oral Reading Test (SORT). 44

Additional Screening Tests. 48

SECTION 5: ACHIEVEMENT TESTS

Adult Basic Education Student Survey (ABESS). 49

Adult Basic Learning Examination (ABLE)

Level I. 55

Level II 61

Level III. 66

Fundamental Achievement Series (FAS). 71

Iowa Tests of Educational Development (ITED). 78

Metropolitan Achievement Tests (MAT). 88

Tests of Adult Basic Education (TABE)

Level E. 98

Level M. 105

Level D. 112

Supplementary Page for TABE. 119

Wide Range Achievement Test (WRAT). 120

Additional Achievement Tests. 124

SECTION 6: DIAGNOSTIC TESTS

Basic Reading Inventory (BRI) 125

Individual Reading Placement Inventory (IRPI) 130

Additional Diagnostic Tests 136

SECTION 7: ESL TESTS

English-Second-Language Placement Test (EPT)

100-200-300. 137

400-500-600. 141

Examination in Structure (EIS). 146

Ilyin Oral Interview (IOI). 150

Oral Placement Test for Adults (SWCEL). 155

Test of Aural Comprehension (TAC) 159

Test of Aural Perception in English for Latin-American Students (TAPE-LAS) 164

Additional ESL Tests 169

APPENDIX

Summary of Test Information 171

Summary of Technical Information 175

References 177

I. IMPROVING EDUCATIONAL PROGRAMS THROUGH THE USE OF STANDARDIZED TESTS

Appropriately selected standardized tests are invaluable tools for improving program effectiveness and efficiency. They are more carefully constructed than teacher-made tests, contain norms (standards) formulated through objective, empirical methods, and provide standard procedures of administration and scoring. Results obtained from standardized tests are, therefore, more accurate gauges for assessing student abilities in particular areas than are those secured through informal testing procedures.

Standardized tests are convenient devices and necessary aids for student placement, evaluating student performance, and measuring program effectiveness. Test scores, however, should not be considered as reflectional of the total individual, but as a measure of performance at a specific point in the student's life - as only one method of measuring performance which adds to the total view of the student.

The end goal of the testing program is to provide a sound basis for making decisions about the total instructional program, its students, the curriculum, etc. Tests should not be used simply for the sake of testing; such indiscriminate use hinders the program. The establishment of the testing program, therefore, necessitates carefully determining what information is needed, how the information is to be used, how the instructional program is organized, the nature of the students, and who will do the testing. Such determinations make it possible to select appropriate tests and derive the greatest potential from test results as decision-making aids.

Types of Useful Tests

Three types of academic tests generally useful in gathering data for making decisions related to the educational program are:

1. Screening tests
2. Achievement tests
3. Diagnostic tests

Screening Tests: Screening tests are short, easy to administer tests, that can be given when the student enters the program in order to make a fast judgment on where to place him, with what work to provide him, or what further tests to give him. Sight word tests that range in difficulty from a non-reading level to a high school level are often used as screening devices in instructional programs. Although some provide grade levels, such grade levels should not be considered as reading achievement levels. Generally, the levels indicate that the words on a certain list are usually taught at that grade level. Sight word tests do not indicate knowledge of word meaning or ability to understand written material; they only indicate the student's ability to recognize and call word names.

Achievement Tests: Achievement tests are comprehensive tests of student performance. They are designed to measure how a student or a group of students compares to the standardization sample (norm) in his knowledge of a particular subject area, such as math, reading, or language skills. Results from achievement tests provide more complete, more reliable, and more detailed information for student placement, assessing program effectiveness, and evaluating student progress than do the more informal short screening tests.

Diagnostic Tests: Occasionally, a student's performance in class or on a group test indicates that more complete and specific information is needed to assist that student. To obtain this information, a diagnostic test may be administered individually.

Diagnostic tests are more detailed than achievement tests. They are designed to discover specific information on an individual's strengths and weakness in a particular skill within a subject area. For example, an achievement test provides an overall score in reading or math, while a diagnostic test provides information on the student's ability to discriminate between various vowel sounds, to divide words into syllables, to use forms of the verb "to be" correctly in conversation, or to divide by two-digit numbers. Because of the importance of accuracy in scorer judgment, diagnostic tests should be given by a trained professional.

One type of information often gained from diagnostic reading tests that is very useful in planning student programs is the relative reading ability of a student on passages of varying difficulty. These levels are the frustration, instructional and independent levels. The frustration level is the level at which the student exhibits great difficulty or inability in word recognition and comprehension, and physical, visible signs of frustration. The instructional level is the level at which he feels comfortable but does not demonstrate complete mastery. This is the level at which further supervised development of his reading ability should take place. Recreational reading, and instructional materials in other subject areas (math, social studies, etc.) should be at the student's independent level, i.e., the level at which he demonstrates little or no problems in word attack skills or in comprehension.

Using Test Results for Decision-Making

The decision-making qualities of test results are useful to teachers and counselors as well as to administrators. Test information integrated with other data, is useful for making decisions in the following areas:

1. Student placement (into a group, curriculum, or material)
2. Identification of special needs
3. Determination of realistic student goals and planning
4. Improvement of personnel understanding
5. Program effectiveness
6. Program evaluation
7. Budget planning
8. Provision of information for outside agencies
9. Improvement of public relations

Student Placement and Instructional Planning: Information from screening, achievement, and/or diagnostic tests is essential in student placement and in planning instructional programs. Placement of students into appropriate groups and instructional materials wherein they will receive instruction commensurate with their level of ability should be partially based on an objective measure of that ability. Placement decisions based on diagnostic tests are, of course, more accurate than those based on achievement tests.

In programs using a highly individualized approach, detailed information provides for a more specific prescription. Information obtained from tests can be objectively discussed with students in planning their specific educational programs. If groups of students are taught, a composite of test data about all the students in the group is useful in deciding which skills should be emphasized in total group instruction, and which should be provided for through individualized instruction in subgroups or in self-directional and programmed materials.

Identification of Special Needs: Review of test performance may indicate that greater information is needed about individual students. An erratic performance on screening or achievement tests suggests that a diagnostic test should be given. Poor results on a test when better results are expected is another indication that further testing is needed.

Occasionally, students may need to be referred to professionally trained persons for psychological or physical testing in order to pinpoint other reasons for inability to perform or progress. Other types of further testing, which may be indicated by performance on a standardized test, are screening tests for vision and for hearing. Data gleaned from these supplementary tests assists in arranging a more conducive instructional environment and provides a sound basis for interpretation of special needs.

Determination of Realistic Student Goals and Planning: Knowledge of test results assists in the counseling of students for immediate and long-range educational and vocational goals. If students know what they are presently capable of doing in relation to their short and long-range goals, they are better able to realize the extent of future training necessary to reach those goals or to modify them. Emphasis here is on what the students can do as compared to what they want to do, and not necessarily on exact grade levels of achievement. (See page 17 for more details on reporting test results to students.)

Improvement of Personnel Understanding: Test results should be shared and discussed with other staff members who are working with the student. This mutual sharing and discussion stimulates an interpretation of the test score in light of all other data staff have observed or collected. Statements and plans formulated should be recorded.

Judging Program Effectiveness: Operation of an instructional program necessitates periodic assessment of the effectiveness of that instruction for individuals and groups. The comparison of entry and retest scores for individuals may be used to decide if a particular instructional plan has been effective for that student. If it has not, then further investigation must be made as to why it has not been effective so that instruction can be changed. The comparison of group entry and retest scores provides the same information about the total group of students, and serves as a basis for instructional decisions for that group. The comparison of test scores between and among groups demonstrates which levels of students are receiving the most effective instruction and which levels of students need a revised curriculum. Finally, a composite of total program gains assists in judging the effectiveness of the program as a whole. Such comparisons of scores imply that the intervals between test and retest are compatible and not widely divergent in length. Scores with an interval of 75 hours should not be compared on an equal basis with those having a 100-hour interval.

Program Evaluation: In conjunction with judging program effectiveness, test scores are valuable in evaluating the design of the program, and its curriculum and emphasis. In combination with enrollment and attendance data and with information on the local population, decisions can be made regarding the appropriateness of the program or segments of the program in meeting its objectives. Such an examination may reveal that the emphasis and curriculum of the program must be changed to accommodate the needs of the people from the target population. In evaluating individual student progress, test information provides an objective basis for moving a student to a higher ABE level, to a High School Equivalency preparatory class or to another program. Also, programs that are carrying on instructional research regarding material or methods need the results of standardized tests so that their findings are acceptable to others.

Budget Planning: Decisions regarding the effectiveness of the program should be reflected in the budget. If it is decided that curriculum and program emphasis need to be changed, then this should be reflected in budgeting for materials and personnel.

Provision of Information for Outside Agencies: Programs which receive students referred through public agencies, industry, or other groups usually have to report the standing or the progress of those students. The use of standardized tests to provide objective information substantiates such reports.

Improvement of Public Relations: Composite test data can be very useful in interpreting the effectiveness of a publicly supported educational program to the public at large. The ability to quote test results when the program is explained to the public lends a professional credibility to program operation.

II. DESIGNING AN EFFECTIVE TESTING PROGRAM

The establishment of an effective testing program depends upon a clear understanding of the purposes for testing, the student population, and program objectives and organization. Based on such knowledge, the criteria for the selection of appropriate test instruments can be finalized. The appropriateness of tests in relation to the information desired should always be the final consideration in test selection.

Technical Considerations

The process of selecting tests should also place great emphasis on a thorough investigation of the following variables inherent in all standardized tests:

1. Validity
2. Standardization sample for norms
3. Reliability
4. Practicality

This information is usually found in the administration manual, the technical manual or the publisher's catalog. Some information, such as validity, norming, and reliability data, is very technical, but essential information can be gleaned without a strong technical background in these areas.

Validity: The essential questions in determining validity are "What does the test measure?," "Is this what you want to measure?" and "Is the content representative of that in your program?" These questions can be answered by examining the publisher's information and the test itself for a comparison with your program's curriculum. Examination of how validity is reported will assist you in determining what the test measures. One or several methods may be used. The following questions will assist you in judging test validity:

1. Does the test exhibit content validity? From what sources were the test items drawn? Are they representative of what is taught to those being tested?
2. Does the test exhibit criterion validity? With what other tests or measures was this test correlated? What is the correlation coefficient? (Correlation coefficients are expressed as numerical values generally ranging from .00 to 1.00, with the more acceptable coefficients being closer to 1.00.) Is this criterion measuring what you want to measure in your students? Criterion validity is essentially saying that this test measures a skill as well as does the criterion measure.
3. Does the test exhibit construct validity? Is the author's rationale for the validity of the test sound? Many tests published for use with adults are revised children's tests. Their subject matter and context are drawn from what is taught in elementary schools, and the criterion measure is usually a test designed for children's use. In some cases, attempts are made to demonstrate that such tests are appropriate for adults through use of adult samples in establishing criterion validity

or through construct validity. In essence, what the authors are saying is that this test measures elementary or junior high school skills, as well as the children's test does with children or adults, but that it is more appropriate because they have used a more adult content or format in which to measure the skills.

Standardization Sample for Norms: So that a test may provide functional information, norms must be reported. The norms represent the performance of a group to which your students will be compared. In looking at the standardization group from which the norms were derived, you should find out if it included students representative of yours in age, sex, educational and geographic background, intelligence, socio-economic level, race, etc.

Most norms reported in test manuals are national norms. To provide a comparison group more appropriate to your students, local norms can be developed. Check the test manual to see if the test publisher will assist programs in developing local norms.

Norms for reading and math, usually reported as grade level norms, are based totally or in part on the performance of children. This means you are comparing the adult's observed ability on the designated material with the ability of children in various grades. Other means of norming such tests which provide useful and accurate information have not yet been widely utilized. Some of these other means of reporting scores are percentile scores, standard scores (as in the GED tests), age-level scores, and stanines.

Reliability: A well constructed test is one that is consistent in its measurement providing ability is held constant. Reliability may be expressed as a standard error of measurement or as a reliability coefficient which usually lies between .00 to 1.00. The standard error of measurement (SE) is used in determining the consistency or reliability of individual scores. When comparing two tests, the one with the lower SE has the greater consistency for individual scores.

Reliability coefficients are most appropriate for determing overall consistency among comparable tests. Three types of consistency measures may be reported: internal (split-half), equivalent form or test-retest (same form) reliability. These various means of reporting reliability yield different reliability coefficients, because each is measuring different aspects of consistency. For example, the lowest coefficient is usually reported in equivalent from reliability and the hightest through test-retest. Yet, equivalent form reliability is a more acceptable measure. Both the type(s) of reliability and the coefficient(s) must be determined. A minimum acceptable reliability coefficient is one that is higher than any other for a similar test instrument reporting the same type of reliability measure.

Practicality: Aside from the technical aspects of judging appropriateness of tests, there are some other questions which must be answered before deciding on a test. Read the manual and publisher's material carefully, and examine the various test components looking for information to ascertain how practical the test is for your program. Information needed includes the details of administration and scoring, the physical properties of the test, the origin, and the cost factors.

While examining a test, note carefully the specific steps involved in administration and scoring. The time aspect, the directions, and the number and use of subtests are important when you are considering how well a test suits your program. This information is also useful in determining what special training is necessary for the staff to give and score the test.

Factors to look for in the physical arrangement of the test include whether or not a test booklet is used and if students answer in booklets or on separate answer sheets. Examine the booklets and answer sheets in order to determine if they are confusing and difficult to follow or if the printing is appropriate in size. Note the different levels and forms of the test series to determine if the range is adequate for the program.

While reading the manual, look for information about the origin of the test and for special adaptations for adult students. Background information on authors and publishers is useful in judging their familiarity with your student population. Also, manuals often contain supplementary explanatory information on using and interpreting test results.

When determining the cost factor of a test, consider the price of its various levels and forms, of the answer sheets, of a scoring service if needed, and of any other materials necessary to administration and scoring. Also, determine which components must be repurchased and the cumulative cost which this entails.

Scheduling Tests and Retests

A testing program should contain a standard policy regarding when tests are to be given. This policy should include when entry tests are administered and when retesting is to take place.

Generally, all students should be tested when they enter the program. Entry tests are especially helpful since so little is known about the student's abilities. If the administration of an achievement test is delayed, the screening test serves to provide minimal information for the student's immediate placement. Achievement testing may take place after the student has attended a few sessions. The delay in achievement testing is often beneficial, for it enables the student to review skills he once knew and to become familiar with his surroundings.

Retesting of each student should take place periodically. The time lapse between test periods, however, is not as important as the consistency of time between test and retest of each student. The time lapse may vary from program to program depending upon the number of instructional hours per week.

Consistency of time periods between tests for every student is important when comparing growth both among individual students and student groups. Ideally, each student should be tested after he has received a specific number of hours of instruction, and at regular intervals thereafter. If an entire class or group is to be tested after a specific number of class sessions or instructional hours have been made available to the group, the actual number of hours of instruction each student has received should be carefully recorded. The progress of a student whose score changes from 5.7 to 6.7 after 95 hours of instruction is not equivalent to that of a student whose scores are identical to his, but who received only 45 hours of instruction during the same time period.

The student should not be given the same test form over and over again, as this tends eventually to produce unreliable results. On the other hand, progress cannot be measured by giving a different achievement test each time, as each test has its own norms, usually not correlated with other tests on the market; therefore, it is best to choose an achievement test that has more than one form, so that while norms are consistent, the student is not exposed to the same questions each time he is tested.

Preparing Students for Standardized Testing

When tests are administered to adult students, care should be taken to provide a positive atmosphere. Most adults, especially the undereducated, are wary of tests and may make a poor showing because of testing conditions. There are several things you can do to alleviate the student's fear:

1. Put the student at ease. Tell him that this is not a pass or fail situation and that you are only asking him to take the test so that you can find out more about his strengths and weaknesses. Impress upon him that this knowledge will help both of you plan a program for him and evaluate the success of that program later on.
2. Before you test the student, you should be satisfied that the test he is taking is appropriate for him. Any doubts or negative feelings you have about the test will be transferred to the student and may negatively affect his performance.
3. If a student is very fearful of a test, provide practice situations which "don't count" on other material so that he feels more comfortable with the test. This is especially useful if IBM type answer sheets are to be used.

4. While administering the test, follow the prescribed directions, but do not build tension in the student by exhibiting nervousness, looking over his shoulder, or telling him to hurry because time is running out.
5. Make sure the student understands the directions before he works on each section.
6. Do not help the student answer the questions. If he asks, tell him you cannot help and to skip that question.
7. Make sure you are prepared. Have all the equipment necessary - test booklet, answer sheets, pencils, extra paper, etc.
8. As much as possible, the test environment should provide good lighting, appropriate furniture, and no distractions.

Interpreting Test Scores

That an appropriate test, designed to measure specific abilities and properly administered, can provide information that is difficult, if not impossible, to obtain in any other way has already been emphasized. It is, however, also necessary to stress that test results must be interpreted in reference to other pertinent information about that student. They are not to be accepted as an absolute, precise measure of student ability. Tests of achievement for adult students are often very close in style to tests given to elementary and junior high school students. The skills measured by such achievement tests do not provide a complete picture of the functional abilities of adults, especially those from other life styles.

The value of test results (especially diagnostic tests) lies very often in their interpretation by a trained and intuitive tester. Individual test scores are samples of student performance on specific skills at a specific point in time, and in comparison to a specific group. In interpreting individual scores, several factors relating to student performance and test construction must be considered. A single test score may be an over or underestimation of a student's ability.

Underestimation of a student's ability may take place if a student panics while being tested or if he is very fearful of tests. In such cases, the student is capable of working on material more difficult than the test score suggests. A delay in administering the entry test may prevent this misrepresentation. Very often reading achievement test scores are indicative of the student's frustration level, i.e., that level at which the student exhibits great difficulty in reading (see page 3). In this case, the student's ability must not be overestimated and the readability level of the material that he receives should be on a lower level than his test score indicates.

To judge the relative accuracy of the test score, compare test performance with the student's ability to read and comprehend reading selections of known difficulty, or to compute arithmetic problems under less stressful circumstances. In reading, it is usually appropriate to select passages one year below the reading ability indicated by the test score, and to advance quickly to the point where the student can read with 70-90% comprehension in a reasonable period of time. In arithmetic, select several problems of the type the student missed and observe his performance.

Another source of misinterpretation of test scores is the acceptance of sub-test scores where reliability data does not warrant such acceptance. Achievement tests often are broken into sub-tests for ease of administering. Sub-test scores are not acceptable measures of student ability in those specific areas, and should not be accepted as such. Poor performance on a particular sub-test indicates only that student ability in that area should be further examined, not that a definite weakness or deficit exists.

Appropriate interpretation of test scores must be in reference to a particular group - the standardization sampling. Since the composition of this group and the method of determining norms differs with each test, comparisons of students with norms across tests to measure progress is a poor practice as different tests will yield different ability levels for the same student. Measurement of progress should be on the same test series to insure compatibility of scores. If more than one test series is to be given, both must be given as pretests and both as post-tests.

Information gained by the staff as they observe the student's performance on prescribed materials should be recorded and considered in the interpretation of test scores. Such information is also valuable in judging student progress between tests. Lack of observable student progress in material and on standardized tests may be an indication that the program planned for him is inappropriate and should be changed. Diagnosis and evaluation of student abilities and progress is an ongoing assessment substantiated by periodic test results.

Reporting Test Results to the Individual

Once the student is tested, it is necessary to tell him the results. At this point, the question of reporting an exact grade level of percentile score to the student is raised. Whether or not this exact score is reported, it is absolutely necessary to report the meaning of the score. To provide an exact score and nothing else is to negate the program's responsibility to the student. A student should receive information on skills tested where he showed ability or inability, and on where this places him in relation to his goals or in relation to daily activities. He should know that the results will be used to determine instruction for him as an individual and later to determine the appropriateness of that instruction. He should know that the test results are not only a measure of his ability, but of the program's ability to fulfill its objectives.

To avoid a misinterpretation of test results by the student, it is desirable and usually possible to avoid the reporting of grade equivalent scores. Several alternatives to reporting grade level scores exist. Since the student knows his position in the program hierarchy, it is possible to report his test results as falling within a range. Another possibility is to report raw scores rather than grade scores. Still a third method is to omit the decimal from a grade equivalent score and to report it as a whole number, e.g., 75 rather than 7.5.

Except in rare instances, reporting and interpretation of test results is best provided by trained personnel. The person inexperienced in test construction is more likely to attribute a precision that does not exist to test scores. Important information available from observation of student performance may be ignored, while the importance of the exact test score is overvalued.

It must be emphasized that the student's score is confidential and should not be shared with other students unless that student chooses to do so. Discussion of test results should, therefore, take place privately where others cannot overhear. The test booklets or profile sheets where the score is written should not be left lying around, but should be kept in a folder and filed unless they are being used.

In conclusion, there are three basic concerns in reporting test scores to the adult student.

1. The interpretation should be set in the frame of reference of the particular learner. Thus, standardized achievement test results should be interpreted in terms of what is known about the learner's aptitude and about his educational and vocational goals.
2. The interpretation should be directed toward positive and constructive action. It should emphasize the assets in an aptitude profile, or it should be oriented toward remedial action when achievement falls below what the student's aptitude leads one to expect. It should point toward realistic and possible educational or vocational goals.
3. The interpretation should be factual and dispassionate, rather than appearing to pass judgment on the individual. Test results and other evidence should be reported truthfully and accurately, but with a friendly and understanding attitude. The flavor should be one of working with the learner to realize common goals.

Criteria of a Testing Program

1. Is the testing program comprehensive in terms of all skill areas?

2. Does the testing program include all adult students in the program?
3. Are the tests given at regular intervals?
4. Are the intervals appropriate to your program needs?
5. Are the results from different tests in the testing program comparable?
6. Do the tests used agree with the objectives and the curriculum of the program?
7. Are the specific tests carefully chosen?
8. Are the tests carefully administered to each group?
9. Are the tests scored accurately?
10. Are the test results interpreted in terms of local as well as national norms?
11. Are the test results quickly reported to teachers and counselors in understandable terms?
12. Are the test results recorded on individual cumulative record forms?
13. Are the test results interpreted to the student in terms of his goals?
14. Is a definite attempt made to relate the test scores to other kinds of information?
15. In addition to the regular testing program, is there provision for special testing as needed?
16. Does the program have an in-service program for educating teachers in the use of test results?
17. Are the test results used to evaluate individual student progress, as well as the total program effectiveness?

Cautions about the use of Tests

1. Don't accept a score as absolute.
2. Don't distribute tests randomly or let students study from them. Test items should provide a representative sampling of the curriculum. When students practice on test items the test score loses its reliability and validity.
3. Don't use test items to teach a skill. You have instructional material for this purpose and should choose material that will prepare the student to take the test later.
4. Keep tests in a secure place where they are not readily accessible to students.
5. Do follow directions.

III. TESTING ENGLISH AS A SECOND LANGUAGE STUDENTS

Many of the considerations mentioned in Parts I and II of this publication are also considerations for English as a Second Language tests. Such tests should exhibit validity, norms, reliability and practicality. The uses of tests in the ESL programs are similar to those in ABE, as is the attention to scheduling tests, preparing students, interpreting scores, and reporting results. What is different about ESL testing is the nature of the information to be obtained and the content of the tests themselves.

The desired informational output of the ESL test is dependent upon the goals of the student and the program. The goals of adult ESL students in New Jersey are usually those of oral rather than written communication. In programs involving such students, the testing or oral rather than written ability is of primary concern. Tests of listening and speaking skills should therefore be used. Most programs also include reading and writing skills in their curriculums, especially when students are academically trained or have some skill in speaking and understanding English. In such programs, knowledge of a student's ability on a reading comprehension test is useful in determining that student's overall comprehension of English. English comprehension tests designed for college students or reading comprehension tests designed for English-speaking ABE students are not appropriate for measuring the reading abilities of most ABE-ESL students, as those tests do not measure the skills the student or the program is seeking to improve. Any reading test chosen should be one designed especially for ESL students. Students whose oral communications skills are non-existent or severely limited should not be asked to take a reading test until their oral skills have improved.

The content and the structure of tests for ABE-ESL students is dependent upon the curriculum of the program. As the test is to be used for initial placement and to measure subsequent improvement, care should be taken to insure that test and program are parallel in content and in method. It is therefore necessary for a program to determine an appropriate number of levels, and to outline the specific listening, speaking, reading, and writing skills to be included at each level, as well as appropriate materials and methods for use in instruction.*

Some standardized tests appropriate for testing ESL students are being published. However, reviews reveal that several of the available tests are presently inadequate for ABE-ESL programs. Most ESL tests, even those designed for ABE-ESL, exhibit only superficial norms, validity, and reliability. (Study the reviews to determine completeness of data for each test.) In many instances data was meager, incomplete and hastily collected. Validity had sometimes been established by correlating the test results with teacher judgment, thus negating the purpose for the test. Populations used to establish the data were narrow - sometimes including only one group or type of student - and small in number. With the exceptions of the Ilyin Oral Interview, and the EPT, and the SWCEL test, the data applied only to ESL students who were seeking academic training.

---

*An ESL curriculum developed in San Francisco for adult students, ESL Master Plan, Phase II, is available from Dr. Steven Morena, Assistant Superintendent, San Francisco Community College District, 33 Gough Street, San Francisco, California. The Ilyin Oral Interview and English-Second-Language Placement Test by Donna Ilyin were designed to place students into this curriculum.

Nevertheless, several ESL test reviews are included in this Guideline. These tests could be used in ESL programs providing they are appropriate for the program curriculum and purposes for testing, and are modified for local use. Generally, the norms will have to be changed to suit the local practices of dividing students into "levels." Local norms (minimal scores for placement into each level/average degree of increase for certain periods of time) will have to be developed over a period of time. The advantages of using one of the published tests is that it is ready-made, will provide some objectivity in measuring placement and improvement, and it can be used as a starting point for revision or development of another test. Some tests, although inadequate in content as general placement and achievement tests are useful for measuring individual skills, such as auditory discrimination or usage.

Because a feasible oral test has not been available, many EST programs utilize an informal interview to judge the oral ability of their students. This type of measurement generally does not yield the objectivity and scores desirable for determining increased ability; therefore, most programs are unable to substantiate claims for success with their ESL students with any identifiable or specific citation of increased ability. In order to use the interview format as a more precise measurement instrument, it is necessary to structure it by incorporating skills included on the various instructional levels in the program, and utilizing the following procedures suggested by Harris:

1. Decide in advance on interview methods and rating standards.
2. Conduct interviews in some quiet place with suitable acoustics.
3. Reserve sufficient time for each interview.
4. Use at least two raters for each candidate.
5. Rate the candidates without reference to other test scores.
6. Record the ratings after the interview.
7. Obtain each candidate's final score by pooling or averaging the two (or more) ratings that have been given him.*

Harris also suggests that oral ability can be measured through "highly structured speech samples...rated according to very specific criteria" or, if other methods of measurement are unavailable, through "paper-and-pencil objective tests of pronunciation..."**

To assist programs in establishing criteria for placement and growth decisions, the chart on the following page is useful. This guide utilizes information concerning oral and graphic skills in determining student placement, and illustrates how information regarding several factors of English communication must be synthesized in judging student competence.

In addition to the published ESL tests, several tests developed by individual programs are available. Although these tests are not standardized, they would be useful for experimental use or as models or to be adapted to particular program uses. A list of these tests are included at the end of the review section in this Guide.

---

*Harris, David P., Testing English as a Second Language. New York: McGraw-Hill Book Company, 1969, pp. 91-92.

**Ibid, p. 83.

ENGLISH PROFICIENCY CHART*

	BEGINNER	HIGH BEGINNER OR LOW INTERMEDIATE	INTERMEDIATE	ADVANCED
LISTENING SKILLS or AURAL COMPREHENSION	Virtually no proficiency	Understands simple questions and statements on familiar topics if spoken very slowly and clearly. Often requires restatement in graphic terms, using gestures and realia.	Understands most questions, statements, and conversations on familiar topics, if spoken clearly at normal speed. Requests occasional restatement.	Understands most informal questions, statements, and conversation at normal speed. Comprehends extended explanations of familiar topics with some difficulty.
SPEAKING SKILLS	Virtually no proficiency	Asks and answers questions on daily personal needs and familiar topics with very limited vocabulary. Makes frequent basic errors in structure and pronunciation.	Converses intelligently in most social situations, but without complete control of structure and pronunciation. Vocabulary is limited.	Participates effectively, sometimes hesitatingly, in social and academic conversations. Makes occasional errors in idiom and structure, often obscuring meaning.
READING AND VOCABULARY	Virtually no proficiency	Reads and understands structures and vocabulary taught orally with written reinforcement in class. Makes constant use of bilingual dictionary.	Reads and understands simple paragraphs involving structures and vocabulary similar to those presented orally in class. Makes constant use of all-English dictionary and frequently refers to bilingual dictionary.	Reads and understands most popular adult materials, e.g. newspapers, magazines, and paperbacks, with regular use of all-English dictionary; and more difficult materials on topics with which he is familiar, with frequent use of bilingual dictionary - all at a labored speed Comprehension difficulties are due to insufficient vocabulary and inability to grasp major ideas.

*Adapted from CIEP English Proficiency Chart
(Commission on Intensive English Programs)

-mor

ENGLISH PROFICIENCY CHART CONTINUED

	BEGINNER	HIGH BEGINNER or LOW INTERMEDIATE	INTERMEDIATE	ADVANCED
WRITING	Virtually no proficiency	Writes simple statements and questions using vocabulary,and structures taught him. Makes frequent errors in spelling and structure, often obscuring meaning.	Writes statements and questions on familiar topics with fair control of basic, but not complex, patterns with frequent obscurity of meaning. Has limited ability to organize a narrative or descriptive paragraph.	Has most sentence structure under fair control within familiar and academic areas, with occasional obscurity of meaning. Under time or test pressures, control weakens. Little understanding of paragraph organization of longer, more formal compositions.

Dolch Basic Sight Word List

General Information:

1.	Title and level	Dolch Basic Sight Word List
2.	Author	Edward Dolch
3.	Publisher	Garrard Publishing Company, Champaign, Illinios
4.	Date of Publication	1942
5.	Cost	100 sheets, $2.50
6.	Time for Administration	About 5 minutes
7.	Number of forms of the test	One
8.	Type of test	Sight word vocabulary
9.	Skill range	Beginning reading (0-3)
10.	Sections, parts, subtests	Two parts

I. Validity:

A.	What does the test measure (content)?	If used individually, ability to immediately pronounce words in isolation. If used as a group test, the ability to identify the written equivalent of a word orally presented by the examiner.
B.	What type of validity does it exhibit?	Content.
	1. From what sources were the test items drawn?	The 220 words which comprise the Basic Sight Word Test were drawn from three word frequency counts compiled from 1926 to 1930. Except for 27 of the 220 words, the basic sight words are those that appeared on all three lists. The types of words on the list include conjunctions, prepositions, pronouns, adverbs, adjectives and verbs. To

substantiate that these words would frequently be encountered by new readers, word counts were taken from texts. The samplings yielded the following information:

Percentage that the Basic Sight Vocabulary is of Running Words in School Textbooks in Four Subjects*

Subject	No. of Series	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Reading	4	70	66	65	61	59	59
Arithmetic	2			62	63	57	57
Geography	2				60	59	54
History	2				57	53	52

A recent study utilizing a frequency count compiled in 1967 on adult reading matter found that 37% of the Dolch 220 words were not among the most frequent 220 words of the new list.

2. With what other tests or instruments was this test correlated?	None.
What is the correlation coefficient?	None.
What does the criterion measure?	None.
3. What is the rationale for the test?	Research completed prior to 1936 indicates that these 220 words make up more than half of the reading matter in school texts. Therefore, knowledge of these words greatly increase a person's ability to read and comprehend. Recognition of these 220 words on sight and in isolation of context is assumed to be

---

*Dolch, Edward W., Teaching Primary Reading. Champaign, Illinois: Garrard Publishing Company, 1950.

	a valid estimation of overall reading ability. To assist in judging relative ability, the author states that normal second-grade readers know about half the words, normal third-grade readers know practically all.*
C. What skills does your program want to measure?	To be answered by program.
D. Examine the test itself.	
1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2. Are the test items well constructed?	To be answered by program.
3. Are they free from ambiguity?	To be answered by program.
E. What have reviewers, critics and users of the test said about the test?	Unavailable.
II. Standardization Sample:	No standardization sample.
A. What was the composition of the persons on whom the test was standardized?	
B. How large was the sample?	
C. From what geographic regions was the sample drawn?	
D. Are local norms available?	

---

*Ibid.

E.	In what terms are the norms reported (grade level, age, etc.)?	
F.	Are students similar to those in your program represented in the sample?	
III.	Reliability:	None reported.
A.	Is a reliability coefficient reported?	
B.	What is this coefficient?	
C.	What type of reliability coefficient was reported?	
D.	What was the composition of the group for whom reliability coefficients were computed?	
E.	Is a standard error of measurement reported?	
F.	What is it?	
IV.	Practicality:	
A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 3 minutes.
C.	Can the subjects be administered at different times?	--
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	None.

F.	Is it group administered?	It can be.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	No.
I.	Are separate answer sheets available?	Each test administration consumes one test sheet.
J.	What is the range for the test series?	No series.
K.	What is the origin of the test?	Research conducted prior to 1936.
L.	How recently has the test been revised?	Never revised.
M.	What were the special adaptations, if any, for adults?	None.
N.	What is the background of the authors and publishers?	Author was quite a prominent person in reading instruction and has authored texts and articles on the subject.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes. Primer size print.
Q.	What are the components to be initially purchased?	Directions, test/answer sheets, scoring key.
R.	What parts of the test are reusable and which must be repurchased?	Answer sheets are consumable.

Harris Graded Word List

General Information:

1.	Title and level	Harris Graded Word List (Queens College Educational Clinic Sample Graded Word Lists)
2.	Author	Albert Harris
3.	Publisher	Originally Queens College Educational Clinic
4.	Date of Publication	Unknown
5.	Cost	Available from the Montclair State College Adult Continuing Education Center without charge
6.	Time for Administration	About 5 minutes
7.	Number of forms of the test	One
8.	Type of test	Sight word vocabulary test
9.	Skill range	0 - 6
10.	Sections, parts, subtests	One

I. Validity:

A.	What does the test measure (content)?	The ability to pronounce on sight words commonly encountered at designated grade levels.
B.	What type of validity does it exhibit?	Content validity.
	1. From what sources were the test items drawn?	The words for preprimer, primer, first, second and third grade lists were drawn from graded vocabulary counts by Clarence Stone. The words for the fourth and fifth grade lists were drawn from vocabulary counts by Donald Durell.
	2. With what other tests or instruments was this test correlated?	None.

	What is the correlation coefficient?	None.
	What does the criterion measure?	None.
	3. What is the rationale for the test?	That ability to call selected words from a graded vocabulary count is a valid indication of reading ability on that grade level.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	The author reports that "these short lists have proven quite useful in clinical practice" with children in the Queens College Educational Clinic.
II.	Standardization Sample:	None reported.
A.	What was the composition of the persons on whom the test was standardized?	
B.	How large was the sample?	
C.	From what geographic regions was the sample drawn?	
D.	Are local norms available?	
E.	In what terms are the norms reported (grade level, age, etc.)?	Grade levels.

F.	Are students similar to those in your program represented in the sample?	
III.	Reliability:	None reported.
A.	Is a reliability coefficient reported?	No.
B.	What is this coefficient?	
C.	What type of reliability coefficient was reported?	
D.	What was the composition of the group for whom reliability coefficients were computed?	
E.	Is a standard error of measurement reported?	
F.	What is it?	
IV.	Practicality:	
A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 2 minutes.
C.	Can the subtests be administered at different times?	No.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Knowledge of the test and scoring procedure.
F.	Is it group administered?	Individually administered.

G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	No manual.
I.	Are separate answer sheets available?	The entire test consists of one page.
J.	What is the range for the test series?	There is no test series.
K.	What is the origin of the test?	A practical screening instrument developed in a clinical situation at Queens College Educational Clinic.
L.	How recently has the test been revised?	No revision.
M.	What were the special adaptations, if any, for adults?	None.
N.	What is the background of the authors and publishers?	The author is a well-known reading specialist. He has written several books and was president of the International Reading Association.
O.	Is the format attractive and easy to follow?	Depends on the copy you are using.
P.	Is the print size appropriate?	Depends on copy you are using. Author recommends primer size type.
Q.	What are the components to be initially purchased?	Copies of the test can be obtained free from the Montclair State College Adult Continuing Education Center.
R.	What parts of the test are reusable and which must be repurchased?	A separate word list must be consumed with each administration.

SRA Reading and Arithmetic Indexes

General Information:

1.	Title and level	SRA Reading and Arithmetic Indexes
2.	Author	SRA Industrial Test Development Staff
3.	Publisher	Science Research Associates 259 East Erie Street Chicago, Illinois 60611
4.	Date of Publication	1968
5.	Cost	Consumable Test Booklets - Arithmetic only or Reading only: $5.95/25; Administrator's Manual $.40
6.	Time for Administration	Approximately 25 minutes for each Index
7.	Number of forms of the test	One
8.	Type of test	Screening
9.	Skill range	1 - 8
10.	Sections, parts, subtests	Reading: 5 contiguous parts: Picture-Word Association Work Decoding Phrase Comprehension Sentence Comprehension Paragraph Comprehension Arithmetic: 4 contiguous parts: Addition and Subtraction of Whole Numbers Multiplication and Division of Whole Numbers Fractions Decimals and Percentages

I. Validity:

A.	What does the test measure (content)?	Reading: 1) Picture-word association: Items require association of a picture with a word; 2) Word decoding of the first letters or the terminal letters of simple words familiar at the fourth grade level or below; 3) Phrase

comprehension: Understanding a simple sentence depends on knowing meaning of preposition; 4) Sentence comprehension: Meaning of complex sentences is dependent on the placement of the modifying words, phrases, or clauses within the sentence; 5) Paragraph comprehension: Paragraphs describe a thing or event or present multiple instructions in sequence.

Arithmetic: 1) Addition computation from single column to multiple columns with carrying; subtraction computation from items involving no borrowing to multiple columns with borrowing and zeros; 2) Multiplication computation from one-digit multiplier to three-digit multiplier and zero; division computation from one-digit divisors to three-digit divisors and remainders; 3) Basic operations involving fractions and fractional mixed numbers; 4) Computations involving decimals and percents.

B.	What type of validity does it exhibit?	Criterion, construct.
	1. From what sources were the test items drawn?	Reading: A pool of items was developed that spanned the five development levels of reading performance. The items in each group were screened by the language department of a Job Corps center to ensure appropriateness of language and general content. Any terminology that might be interpreted differently by members of various cultures was eliminated and questionable items were rewritten or deleted. Arithmetic: Basic arithmetic computations were chosen and grouped to measure basic skill areas required for four types of jobs.
	2. With what other tests or instruments was this test correlated?	SRA Pictorial Reasoning Test (PRT). In order to determine the degree to which scores on RAI are indepen-

	dent of general ability, an attempt was made to minimize correlations between items and general ability as measured by the PRT. The maximum correlation accepted between item and PRT score was .40.
What is the correlation coefficient?	Reading Index: .229. Arithmetic Index: .227. Arithmetic Index to Reading Index: .463.
What does the criterion test measure?	PRT is a measure of reasoning ability containing no reading or arithmetic items.
3. What is the rationale for the test?	RAI is designed for use with applicants for entry-level jobs and special training programs, where the basic skills of the applicants are often too low to be reliably evaluated by typical selection tests. Most other tests designed for adults assume proficiency above that possessed by many applicants, and they assume levels of proficiency above those actually required by the jobs for which they are used. RAI evaluates an individual's reading and arithmetic proficiency in terms of what skills are actually required on various types of jobs.
C. What skills does your program want to measure?	To be answered by program.
D. Examine the test itself.	
1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2. Are the test items well constructed?	To be answered by program.
3. Are they free from ambiguity?	To be answered by program.

E.	What have reviewers and critics and users of the test said about the test?	In Buro's 7th Mental Measurements Yearbook, Dorothy C. Adkins makes the following comments: The items seem to have been adequately edited, pretested, and analyzed with respect to internal criteria and difficulty. The practice of determining whether students are below potential by using RAI along with a test of general ability, as described in the manual, is questionable. The burden of proof of validity rests on the user. The tests may be useful in certain appropriate situations.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	Index levels were not determined by collecting a sample of student scores, but by defining specific tasks required at various levels of jobs.
B.	How large was the sample?	Not applicable.
C.	From what geographic regions was the sample drawn?	Not applicable.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported (grade level, age, etc.)?	5 index levels of reading tasks required by various jobs, and 4 index levels of arithmetic proficiency required by jobs.
F.	Are students similar to those in your program represented in the sample?	Not applicable.

III. Reliability:

A.	Is a reliability coefficient reported?	No.
B.	What is this coefficient?	Reading: .87; Arithmetic: .91 and .87.

C.	What type of reliability coefficient was reported?	Split-half.
D.	What was the composition of the group for whom reliability coefficients were computed?	Reading: 87 males and females in a combination of on-job training and basic education in Chicago, ages ranging from 17 to 30 years. Arithmetic: 1) 57 persons from Chicago program; 2) 419 students in Colorado and South Carolina Adult Education programs.
E.	Is a standard error of measurement reported?	No.
F.	What is it?	None.

Another way to view the reliability of a test scored in the manner that the RAI is scored is to consider the consistency and meaningfulness of classification. Using the "passing" score for each level of each index, the trainees in the Chicago group were classified on proficiency. Of the 87 trainees taking the Reading Index, 4 (4.6 percent) were classified inconsistently (failed at one level but passed at a higher one) and of the 57 taking the Arithmetic Index, 4 (7 percent) were classified inconsistently. These data were analyzed by computing the Raju-Guttman Index of Homogeneity, a statistic that measures the degree to which it can be assumed that if a subject passes a certain item, he will also have passed all previous items. The Index of Homogeneity for reading is .93; for arithmetic, .95.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 2-3 minutes for each index.
C.	Can the subtests be administered at different times?	Yes.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.

E.	What qualifications are needed to administer the test?	None.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	No. Responses are recorded on scoring grids inside the sealed pages of the test booklet.
J.	What is the range for the test series?	Grade levels 1-8.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	Has not been revised since published in 1968.
M.	What were the special adaptations, if any, for adults?	Designed especially for adults.
N.	[illegible] is the background [illegible] authors and pub[illegible]s?	Science Research Associates is one of the leading publishers of tests and educational materials.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets and Administrator's Manual.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets are consumable.

Slosson Oral Reading Test

General Information:

1.	Title and level	Slosson Oral Reading Test
2.	Author	Richard L. Slosson
3.	Publisher	Slosson Ed. Publications, 140 Pine St., East Aurora, New York
4.	Date of Publication	1963
5.	Cost	$.75 for one pad of 20 sheets
6.	Time for Administration	About 10 minutes
7.	Number of forms of the test	One
8.	Type of test	Sight word vocabulary/screening
9.	Skill range	0.0 - High School (general category)
10.	Sections, parts, subtests	One only

I. Validity:

A.	What does the test measure (content)?	Ability to pronounce words of varying difficulty.
B.	What type of validity does it exhibit?	Content, criterion.
	1. From what sources were the test items drawn?	From standardized school readers.
	2. With what other tests or instruments was this test correlated?	Standard Oral Reading Paragraphs by William S. Gray.
	What is the correlation coefficient?	.96.
	What does the criterion test measure?	Oral reading of paragraphs of ordered difficulty, speed, comprehension.
	3. What is the rationale for the test?	Information unavailable.

C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Unavailable.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	Grade school children.
B.	How large was the sample?	Unknown.
C.	From what geographic regions was the sample drawn?	Unknown.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported (grade level, age, etc.)?	Grade level.
F.	Are students similar to those in your program represented in the sample?	Unknown.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.

B.	What is this coefficient?	.99.
C.	What type of reliability coefficient was reported?	Test-retest.
D.	What was the composition of the group for whom reliability coefficients were computed?	Unknown.
E.	Is a standard error of measurement reported?	No.
F.	What is it?	None.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 3-5 minutes.
C.	Can the subtests be administered at different times?	No.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	No special qualifications. Administration should be thorough and familiar with directions.
F.	Is it group administered?	No. Individually.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	No.
I.	Are separate answer sheets available?	Test itself is only one page which must be consumed.

J.	What is the range for the test series?	0.0 - H.S.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	Published in 1963.
M.	What were the special adaptations, if any, for adults?	None.
N.	What is the background of the authors and publishers?	Author of Slosson Intelligence Test.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Package of test/answer sheets.
R.	What parts of the test are reusable and which must be repurchased?	One test page is consumed with each student.

Additional Screening Tests that you may wish to obtain and review:

Published Tests

Gilmore Oral Reading Test

Harcourt Brace Jovanovich
757 Third Avenue
New York, New York 10017

Gray Oral Reading Tests, 1963 Edition

Bobbs-Merril Company, Inc.
Test Division
4300 West 62nd St.
Indianapolis, Ind. 46206

Los Angeles Diagnostic Tests: Fundamentals of Arithmetic

California Test Bureau
Del Monte Research Park
Monterey, Calif. 93940

Reading Progress Scale

Revrac Publications
1525 Oak Drive
Silver Springs, Md.20910

Unpublished Tests

Cloze Tests for Adults by Joan Fischer

Available from:

Adult Education Resource Center
Montclair State College
Upper Montclair, N.J. 07043

Informal Reading Inventory by Joan Fischer

Available from:

Adult Education Resource Center
Montclair State College
Upper Montclair, N.J. 07043

VIP Math Placement Test by Jane F. Flaherty

Available from:

Adult Education Resource Center
Newark State College
Union, N.J. 07083

Adult Basic Education Student Survey

General Information:

1.	Title and level	Adult Basic Education Student Survey
2.	Author	Elvin Rasof and Monroe C. Neff
3.	Publisher	Educational Opportunities Division Follett Publishing Company 1010 W. Washington Blvd. Chicago, Illinois 60607
4.	Date of Publication	1966
5.	Cost	Materials for instructor and pre and post-testing of 20 students: $23.55; Pkg. of 20 booklets, Pts. 1 & 2 or 3 & 4: $6.60; Pkg. of 100 answer sheets: $9.90; Answer templates Pts. 1-4: $1.50; manual: $1.50
6.	Time for administration	Test is untimed. Manual recommends that from 2-4 sessions should be allowed for administration. It suggests that all students should be able to complete each part within an hour and a quarter
7.	Number of forms of the test	Two; A and B
8.	Type of test	Achievement
9.	Skill range	0 - 8
10.	Sections, parts, subtests	Four: 1) Reading Comprehension; 2) Word Recognition; 3) Arithmetic Computation; and 4) Arithmetic Problems

I. Validity:

A.	What does the test measure (content)?	Part 1: Primarily literal comprehension of silent reading as measured by multiple choice questions. Part 2: Word meaning in context. Students must complete a sentence with the appropriate word out of a choice of 4. Part 3: Ability to add, subtract, divide and multiply whole numbers,

decimals, some fractions and measures. Ability to recognize the meaning of +, −, ÷, x.

Part 4: Ability to read sentences and recognize appropriate answers expressed in math symbols. Requires knowledge of time relationships, money relationships and symbols, linear measures, and arithmetic operations.

About 1/3 of the items are practical problems. Most test straight computation or numerical concept. To correct for guessing, students are penalized for incorrect answers. This helps to insure accuracy of scores.

B.	What type of validity does it exhibit?	Criterion (Predictive validity) for Reading Sections only. Exhibits no validity for other uses besides placement into vocational categories cited in study.
	1. From what sources were the test items drawn?	Unknown.
	2. With what other tests or instruments was this test correlated?	With exhibited ability to complete skills training in several manpower-training classes.
	What is the correlation coefficient?	Coefficients range from .63 to .94. Most of the coefficients fall in the .70's. This indicates that a specific score on the survey is an accurate predictor of ability to complete a skills training class in a specific topic. (Grade Equivalent scores from other tests could not be used for this prediction since grade equivalents from other tests have not been correlated with those on the ABE Student Survey.)
	What does the criterion test measure?	This test was not correlated with another achievement test to establish the validity of scores for purposes other than prediction.
3.	What is the rationale for the test?	That tests designed for children were inadequate for placing ABE

		students, for measuring their growth and for preparation for life.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	Dorothy C. Atkins thinks not. See review in Buro's *7th Mental Measurements Yearbook*.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Reviews from Buro's *7th Mental Measurements Yearbook*, by Dorothy C. Atkins and A.N. Hieronymous, indicate poor test construction, vague standardization procedures, and questionable validity and reliability. Since no criterion test was used to establish norms, grade-equivalents appear too arbitrary. Hieronymous concludes: "This bettery might be useful in administering and conducting Adult Basic Education programs, but it meets few of even the minimum standards for test publication."
II.	*Standardization Sample:*	The exact procedures used to determine grade equivalents for raw scores were not described. It is unclear how grade equivalents were determined from a sampling composed entirely of out-of-school adults.
A.	What was the composition of persons on whom the test was standardized?	Negro, Caucasian, and Mexican-American ABE students in Detroit and No. Carolina. The students were of varied ages and came from urban and rural areas. (Unclear whether norms were revised after use with students from No. Carolina or whether different populations were used to arrive at norms for each form.)

B.	How large was the sample?	Total exact figure is not stated.
C.	From what geographic regions was the sample drawn?	Detroit. Later used with non-urban students in No. Carolina.
D.	Are local norms available?	Unknown.
E.	In what terms are the norms reported (grade level, age, etc.)?	Raw scores, grade equivalents, percentiles, and standard scores.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is this coefficient?	Ranged from .73 to .96 depending upon time interval between pre and post-testing.
C.	What type of reliability coefficient was reported?	Equivalent form.
D.	What was the composition of the groups for whom reliability coefficients were computed?	10 percent of the norming population.
E.	Is a standard error of measurement reported?	No.
F.	What is it?	None.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 3 minutes for each part.
C.	Can the subtests be administered at different times?	Yes.

D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Knowledge of test procedure and scoring.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Yes. Sample exercises clarify directions.
H.	Does the manual contain guides for using and interpreting the scores?	Yes. In terms of predictive validity.
I.	Are separate answer sheets avaiable?	Yes. IBM type.
J.	What is the range for the test series?	0 - 8.
K.	What is the origin of the test?	Authors felt the need for a test designed especially for adults. Test constructed at Skill Center in Detroit.
L.	How recently has the test been revised?	Form A published in 1966; Form B in 1967.
M.	What were the special adaptations, if any, for adults?	Test designed especially for the adult illiterate below 4th grade-level.
N.	What is the background of the authors and publishers?	Monroe Neff has been active in Adult Education as the Director of the Division of Continuing Education in New York and as past-president of NAPCAE. Dr. Elvin Rasof has served as Director of the In-School Youth Work-Training Program in Detroit. Follett Publishing Co. publishes many materials for adults.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Use of all capital letters for beginning level reading sections makes

		those sections harder to read as the students are unable to use configuration clues to decipher words.
Q.	What are the components to be initially purchased?	Manual, multiple copies of test booklets for Forms A & B. Two booklets per form. Answer sheets. Scoring templates.
R.	What parts of the test are reusable and which must be repurchased?	If answers are recorded in the booklets, they cannot be reused. If IBM Answer Sheets are used, booklets are reusable.

Adult Basic Learning Examination, Level I

General Information:

1.	Title and level	Adult Basic Learning Examination (ABLE), Level I
2.	Author	Bjorn Karlsen, Richard Madden, and Eric F. Gardner
3.	Publisher	Harcourt, Brace, and World, Inc. 757 3rd Ave., New York, N.Y. 10017
4.	Date of Publication	1967
5.	Cost	$17.50 pkg/35 test booklets, Administrator's Handbook & Answer Key included
6.	Time for Administration	Work time is approximately 2 hrs., 10 min. Additional time must be allowed for distribution of materials, giving directions, etc.
7.	Number of forms of the test	Two - Form A and Form B
8.	Type of test	Battery of Achievement Tests
9.	Skill range	Grade levels 1-4
10.	Sections, parts, subtests	Test 1: Vocabulary Test 2: Reading Test 3: Spelling Test 4: Arithmetic Part A: Computation Part B: Problem Solving

I. Validity:

A.	What does the test measure (content)?	Spelling (dictated). Vocabulary (sentences and completion choices dictated - no reading involved). Reading Comprehension (choosing one of three choices for a missing word in a sentence or paragraph. No time limit). Arithmetic Computation (operations with whole numbers. No time limit).

		Arithmetic Problem Solving (dictated - testee chooses from possible answers given in the test booklet. Mostly one-step whole number and money computations and "common sense" problems. No reading involved. No time limit).
B.	What type of validity does it exhibit?	Criterion, construct.
	1. From what sources were the test items drawn?	Adult educators.
	2. With what other tests or instruments was this test correlated?	Stanford Achievement Test.
	What is the correlation coefficient?	Vocabulary: .60; Reading: .62; Spelling: .76; Computation: .71; Problem Solving: .73.
	What does the criterion test measure?	Word Meaning, Paragraph Meaning, Spelling, Arithmetic Computation, Arithmetic Concepts, and Arithmetic Applications.
	3. What is the rationale for the test?	The examination was developed in consultation with a variety of people working with the many facets of adult education and was written according to the following general guidelines: 1) Content which is adult oriented; 2) Measurement of achievement as low as first grade; 3) Format which does not appear childish and is easy to follow; 4) Coverage of only the fundamental areas; efficient in use of time; 5) Simplicity of administration for teachers with relatively limited experience in testing.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate	

	for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics, and users of the test said about the test?	Three reviews of ABLE, Levels I and II, by A.N. Hieronymous, Edward B. Fry, and James W. Hall, are included in Buro's 7th Mental Measurements Yearbook. Some weaknesses they point out are: 1) The type of item used in the reading test excludes such comprehension skills as generalization, getting the main idea, and interpretation of author's intention (Hieronymous); 2) ABLE I should be used for Grade Levels 3-8 (Fry); 3) No test is provided for assessing capitalization, punctuation, and usage skills (Hieronymous). Hieronymous concludes, however, "This is a well-conceived, well-constructed battery which should serve very well the purposes for which it was intended. Because it was developed specifically for use with adult groups, because of supporting data presented in a professional manner, and because of its generally high quality, the ABLE is recommended for use over the Adult Basic Education Student Survey or the Tests of Adult Basic Education" Comments by users of ABLE I includes: 1.The reading test tends to yield inflated scores. 2.It is particularly useful for testing persons with limited literacy skills, since the vocabulary and arithmetic problems tests require no reading of the student.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	Children from Grades 2,3,4 and 5, drawn from four school systems in four states, providing a wide range of ability, but with a preponderance of pupils below average in academic achievement.
B.	How large was the sample?	4000: 1000 from each grade.
C.	From what geographic regions was the sample drawn?	From four states (not specified).
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported (grade level, age, etc.)?	Number correct converted to grade score. In addition, stanines are given for No. Carolina Prison Inmates, Connecticut ABE students and Norfolk, Va. ABE students, in a supplementary report.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported.	Yes, for 7 groups (see D below). Given for Vocabulary, Reading, Spelling, Computation, Problem Solving, and Total Arithmetic - None given for the total battery.
B.	What is this coefficient?	For 3rd and 4th graders they range from .73 to .95; median .91. For adult groups, they range from .84 to .98; median .93.
C.	What type of reliability coefficient was reported?	Split-half.
D.	What was the composition of the group for whom reliability coefficients were computed?	Seven groups: 3rd graders, 4th graders, Job Corps Enrollees, Hartford-New Haven ABE Students, No. Carolina Prison Inmates, Connecticut ABE Students, Norfolk, Va. ABE Students.

E.	Is a standard error of measurement reported?	Yes, for each subtest, for the Job Corps Enrollees and the Hartford-New Haven ABE Students.
F.	What is it?	Range from 1.5 to 3.1; median 2.4.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 20-30 minutes.
C.	Can the subtests be administered at different times?	Yes.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Paraphrased from the Handbook: Insofar as it is possible, all tests should be administered by person familiar to the students in order to guarantee a relaxed, non-threatening atmosphere for testing and to ensure that the examiner's enunciation, regional accent and speech patterns will be familiar to the students, especially in the dictated tests.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	The Handbook suggests that additional proctors be present to give individual assistance. Since the tests are not timed, students who are slow in understanding directions are not penalized.
H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	No.
J.	What is the range for the test series?	Grade levels 1-12.

K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	First published in 1967 - No revisions have been made.
M.	What were the special adaptations, if any, for adults?	The test was designed specifically for undereducated adults.
N.	What is the background of the authors and publishers?	The authors are psychologists, and have experience in developing other widely-used achievement tests. Drs. Karlsen and Gardner have backgrounds in ABE. Harcourt, Brace, and World is a reputable test publisher, having produced a number of widely-used tests.
O.	Is the format attractive and easy to follow?	Fry says yes in Buro's 7th Mental Measurement Yearbook.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets, Administrator's Handbook, answer key.
R.	What parts of the test are reusuable and which must be repurchased?	Test booklets are consumable.

Adult Basic Learning Examination, Level II

General Information:

1.	Title and level	Adult Basic Learning Examination (ABLE), Level II
2.	Author	Bjorn Karlsen, Richard Madden, and Eric F. Gardner
3.	Publisher	Harcourt, Brace & World, Inc., 757 3rd Ave., New York, N.Y. 10017
4.	Date of publication	1967
5.	Cost	$18.00 pkg/35 test booklets, Administrator's Handbook & Answer Key included
6.	Time for administration	Work time is approximately 2 hrs., 15 min. Additional time must be allowed for distribution of materials, giving directions, etc.
7.	Number of forms of the test	Two - Form A and Form B
8.	Type of test	Battery of Achievement Tests
9.	Skill range	Grade levels 5-8
10.	Sections, parts, subtests	Test 1: Vocabulary Test 2: Reading Test 3: Spelling Test 4: Arithmetic Part A: Computation Part B: Problem Solving

I. Validity:

A.	What does the test measure (content)?	Vocabulary (sentences and completion choices dictated - no reading involved). Reading Comprehension (choosing one of three choices for a missing word in a sentence or paragraph. No time limit). Spelling (dictated). Arithmetic Computation (primarily computations with whole numbers and fractions, with some percentage, equations, and negative numbers).

		Arithmetic Problem Solving (computation with whole numbers, money, fractions, decimals, percentage, and graph reading).
B.	What type of validity does it exhibit?	Criterion, construct.
	1. From what sources were the test items drawn?	Adult educators.
	2. With what other tests or instruments was this test correlated?	Stanford Achievement Test.
	What is the correlation coefficient?	Vocabulary: .65; Reading: .76; Spelling: .82; Computation: .75; Problem Solving: .66.
	What does the criterion test measure?	Word Meaning, Paragraph Meaning, Spelling, Arithmetic Computation, Arithmetic Concepts, and Arithmetic Applications.
	3. What is the rationale for the test?	The examination was developed in consultation with a variety of people working with the many facets of adult education and was written according to the following general guidelines: 1) Content which is adult oriented; 2) Measurement of achievement as low as first grade; 3) Format which does not appear childish and is easy to follow; 4) Coverage of only the fundamental areas; efficient in use of time; 5) Simplicity of administration for teachers with relatively limited experience in testing.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.

2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	See Buro's reviews and comments under ABLE I, page 56.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	Children from Grades 5,6, and 7, drawn from four school systems in four states, providing a wide range of ability, but with a preponderance of pupils below average in academic achievement.
B.	How large was the sample?	3000: 1000 from each grade.
C.	From what geographic regions was the sample drawn?	From four states (not specified).
D.	Are local norms available?	
E.	In what terms are the norms reported (grade level, age, etc.)?	Number correct converted to grade score. In addition, stanines are given for No. Carolina Prison Inmates, Connecticut ABE Students, and Norfolk, Va. ABE Students, in a supplementary report.
F.	Are students similar to those in your program represented in the sample?	

III. Reliability:

A.	Is a reliability coefficient reported?	Yes, for 7 groups. (See D below.) Given for Vocabulary, Reading, Spelling, Computation, Problem Solving, and Total Arithmetic. No coefficients are given for the total battery.
B.	What is this coefficient?	For 6th and 7th graders, they range from .60 to .96; median .89. For adult groups, they range from .82 to .97; median .91.

C.	What type of reliability coefficient was reported?	Split-half.
D.	What was the composition of the group for whom reliability coefficients were computed?	Seven groups: 6th Graders, 7th Graders, Job Corps Enrollees, Hartford-New Haven ABE Students, No. Carolina Prison Inmates, Connecticut ABE Students, Norfolk, Va. ABE Students.
E.	Is a standard error of measurement reported?	Yes, for each subtest, for the Job Corps Enrollees and the Hartford-New Haven ABE Students.
F.	What is it?	Range from 1.6 to 3.0; median 2.3.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 20-30 minutes?
C.	Can the subtests be administered at different times?	Yes.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Paraphrased from the Handbook: Insofar as it is possible, all tests should be administered by a person familiar to the students in order to guarantee a relaxed, non-threatening atmosphere for testing and to ensure that the examiner's enunciation, regional accent, and speech patterns will be familiar to the students, especially in the dictated tests.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	The Handbook suggests that additional proctors be present to give individual assistance. Since the tests are not timed, students who are slow in understanding directions are not penalized.

H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	No.
J.	What is the range for the test series?	Grade levels 1-12.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	First published in 1967 - No revisions have been made.
M.	What were the special adaptations, if any, for adults?	The test was designed specifically for undereducated adults.
N.	What is the background of the authors and publishers?	The authors are psychologists, and have experience in developing other widely-used achievement tests. Drs. Karlsen and Gardner have backgrounds in ABE. Harcourt, Brace & World is a reputable test publisher, having produced a number of widely-used tests.
O.	Is the format attractive and easy to follow?	Fry says yes in Buro's 7th Mental Measurement Yearbook.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets, Administrator's Handbook, answer keys.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets are consumable.

Adult Basic Learning Examination, Level III

General Information:

1.	Title and level	Adult Basic Learning Examination (ABLE), Level III
2.	Author	Bjorn Karlsen, Richard Madden and Eric F. Gardner
3.	Publisher	Harcourt, Brace & World, Inc., 757 3rd Ave., New York, N.Y. 10017
4.	Date of publication	1971
5.	Cost	Test booklets, Administrator's Handbook included: $17.50; 35 IBM 1230 Answer Sheets: $2.80; IBM 1230 Answer Key: 80¢ each
6.	Time for administration	Work time is approximately 3 hrs., 30 min. Additional time must be allowed for distribution of materials, giving directions, etc.
7.	Number of forms of the test	Two - Form A and Form B
8.	Type of test	Battery of Achievement Tests
9.	Skill range	Grade levels 9-12
10.	Sections, parts, subtests	Test 1: Vocabulary Test 2: Reading Test 3: Spelling Test 4: Arithmetic Part A: Computation Part B: Problem Solving

I. Validity:

A.	What does the test measure (content)?	Vocabulary (sentences dictated - testee chooses from four words given in test booklet. No reading necessary). Spelling (indicating incorrectly spelled words - multiple choice). Reading Comprehension (comprehension questions on paragraphs and a facsimile of a newspaper's front page. The multiple choice items test ability to comprehend explicitly

	stated facts, make inferences, draw conclusions and recall specific information). *Arithmetic Comprehension* (primarily operations with whole numbers, fractions, decimals, measures, and percents; also includes ratios, negative numbers, series, application of number properties, square roots, algebraic equations, and operations with exponential numbers. Multiple choice). *Arithmetic Problem Solving* (primarily involving arithmetic computation, reading of graphs, and measures; also includes a variety of other concepts and skills. Multiple choice).
B. What type of validity does it exhibit?	Criterion, construct.
1. From what sources were the test items drawn?	Adult educators.
2. With what other tests or instruments was this test correlated?	Three subtests of Stanford Achievement Test, Advanced Battery, Form W.
What is the correlation coefficient?	Correlations between ABLE III Subtests and Stanford Subtests for each of seven groups are given: some for ABLE Form A, and some for ABLE Form B. Coefficients for similar subtests range from .53 to .90 with most of them over .75.
What does the criterion test measure?	Spelling, Reading, and Numerical Competence.
3. What is the rationale for the test?	The examination was developed in consultation with a variety of people working with the many facets of adult education and was written according to the following general guidelines: 1) Content which is adult oriented; 2) Measurement of achievement from first grade through high school level; 3) Format which does not appear childish and is easy to follow; 4) Coverage of only the fundamental areas; efficient in use of time; 5) Simplicity of administration for

		teachers with relatively limited experience in testing.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Review of ABLE III is not included in Buro's 7th Mental Measurement Yearbook.
II.	Standardization Sample:	
A.	What was the composition of the persons on whom the test was standardized?	High school students in Grades 10 and 12.
B.	How large was the sample?	1050.
C.	From what geographic regions was the sample drawn?	Communities in Florida, Massachusetts, and Minnesota, chosen in order to provide a sample with a wide range of ability, socio-economic background and geographical location, participated in this phase of the research.
D.	Are local norms available?	To be answered by program?
E.	In what terms are the norms reported (grade level, age, etc.)?	Number correct converted to grade score. In addition, percentiles and stanines for Grades 9 through 12, and for each adult group given.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes, for Vocabulary, Spelling, Reading, Computation, Problem Solving, and Total Arithmetic for each of the 7 groups. None given for the total battery.
B.	What is this coefficient?	Range from .75 to .96, with most of them over .86.
C.	What type of reliability coefficient was reported?	Equivalent form.
D.	What was the composition of the group for whom reliability coefficients were computed?	Besides the groups of 10th and 12th Graders, 5 adult groups totaling 775 persons were included. They consisted of High School Equivalency candidates, vocational training students, and W.I.N. program students in Kansas, Louisiana, Tennessee, New Jersey and Newfoundland.
E.	Is a standard error of measurement reported?	Yes, in raw scores for each subtest, for each group.
F.	What is it?	Range from 2.7 to 4.4; median 3.3.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	About 10 minutes, if IBM answer sheets are used.
C.	Can the subtests be administered at different times?	Yes.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Ability to follow directions exactly as given in the Administrator's Handbook.
F.	Is it group administered?	Yes.

G.	Are the directions easily followed by students?	If directions are given properly, few students on this level will have difficulty.
H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	Yes.
J.	What is the range for the test series?	Grade levels 1-12.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	First published in 1971.
M.	What were the special adaptations, if any, for adults?	The test was designed specifically for undereducated adults.
N.	What is the background of the authors and publishers?	The authors are psychologists, and have experience in developing other widely-used achievement tests. Drs. Karlsen and Gardner have backgrounds in ABE. Harcourt, Brace & World is a reputable test publisher, having produced a number of widely-used tests.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets, Administrator's Handbook, answer keys.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets are consumable.

Fundamental Achievement Series

General Information:

1.	Title and level	Fundamental Achievement Series
2.	Author	George K. Bennett and Jerome E. Doppelt
3.	Publisher	The Psychological Corp., 304 E. 45th St., New York, N.Y. 10017
4.	Date of publication	1968
5.	Cost	Set of 25 complete tests, tapes, scoring keys, manual: $21.00; Specimen Set (no tapes): $1.00
6.	Time for administration	1 hour
7.	Number of forms of the test	Two: A and B
8.	Type of test	Achievement
9.	Skill range	Basic literacy to 8th Grade level
10.	Sections, parts, subtests	Test 1: Verbal; Test 2: Numerical

I. Validity:

A.	What does the test measure (content)?	Verbal: Several types of test items are utilized to measure practical verbal ability. These include: ability to identify the correct bus to a designated destination; ability to derive meaning from signs; ability to identify price of menu items; to use an apartment house directory and others. Answers involve factual as well as inferential skill. Easy and difficult items are throughout the test. Numerical: Test items require the practical use of math skills Items include the ability to understand time and money relationships; to translate words into number symbols; to comprehend charts and tables; to compute math problems; to understand word problems.

B.	What type of validity does it exhibit?	Criterion, construct. Only validity information that pertains to both Forms A and B is cited. Additional information on the validity of Form A is available in the manual.
	1. From what sources were the test items drawn?	Items similar to those encountered by adults in their everyday life were used on the test. These include menus, schedules of varying sorts, telephone directories, radio announcements, sales slips, checks, charts and tables.
	2. With what other tests or instruments was this test correlated?	1) Differential Aptitude Tests (Form A) Numerical Ability. 2) Gates Reading Survey (Form 3) Comprehension. 3) Personnel Tests for Industry (Form S) Group A, Group B. 4) Stanford Achievement Test (Forms W and X) Paragraph Meaning, Arithmetic Computations. All of the above studies utilized adults from anti-poverty training programs in New York and another eastern city. Correlation coefficients were also computed for performance criteria which included supervisors' ratings, teachers' ratings, counselor ratings, and student attendance from several different programs.
	What is the correlation coefficient?	Only correlations with total scores will be reported here. Further detailed information is available in the 1970 manual supplement.

Group Designated by number	Correlation with FAS Total Score (V & N) except where otherwise indicated.
1	.59
2	.53
3a	.45
3b	.67
4PM	.64 (with verbal section only)
4AC	.57 (with numerical section only)

Other correlations ranged from .04 to .63.

What does the criterion test measure?	The Differential Aptitude Test yields scores for verbal reasoning, numerical ability and a combined score. The Gates Reading Survey provides a grade level score for comprehension. For information on the Stanford Achievement, see review on IRPI and ABLE in this publication.
3. What is the rationale for the test?	That a test composed of items thought to be familiar to both advantaged and disadvantaged persons would yield useful information for making decisions relative to employment for disadvantaged persons. Validity studies conducted have sought to establish test validity for this use.
C. What skills does your program measure?	To be answered by program.
D. Examine the test itself.	
1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2. Are the test items well constructed?	To be answered by program.
3. Are they free from ambiguity?	To be answered by program.
E. What have reviewers, critics and users of the test said about the test?	Buros' reviewers (7th Mental Measurements Yearbook) cited the lack of an item analysis of the test items to determine which are difficult items, questioned the small number of persons in the reliability studies, and criticized the lack of validity information. Although the reviewers criticized the Psychological Corp. for the lack of its usual psychometric information, they observed that this test showed promise as a tool for educational and vocational placement of the disadvantaged. Reviews were done by Norman Fredericksen and Lewis R. Aiken.

II.	Standardization Sample:	In standardizing this test, several specific groups were utilized. The findings were not consolidated. This procedure allows the examiner to relate each person's ability to that of a specified group, and to make decisions based on this relationship.
A.	What was the composition of the persons on whom the test was standardized?	For Form A, standardization information was computed for Negro and White students in a southern city school system for grades 6,8,10 and 12; for Negro and White students in a northern city school system for grades 6,8 and 10; and for applicants and employees in a southern food processing plant. For Forms A and B, standardization information was computed for steel plant production employees in a western city, for service employees at a southern hospital, for employees in private anti-poverty training programs in two cities, and for trainees in public anti-poverty programs in several locations.
B.	How large was the sample?	Each sample was composed of at least 100 persons or more. For Forms A and B, the total number for verbal was 1175, for numerical 1181, and for combined scores 1174.
C.	From what geographic regions was the sample drawn?	From western, northern, southern and eastern areas.
D.	Are local norms available?	The manner of reporting the norms in small groups is equivalent to local norms.
E.	In what terms are the norms reported (grade level, age, etc.)?	Percentiles, raw scores.
F.	Are students similar to those in your program represented in the sample?	

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.

B.	What is this coefficient?	Several are reported for each form. For Form A, split-half reliability was reported for six different groups. The correlations for total FAS scores ranged from .84 to .97 with most in the .90's. Correlations for verbal and numerical sections alone were lower. Four studies were also done to assess test-retest reliability. The coefficients for total scores were .96,.94,.72 and .92. Split-half reliability for Form B was computed for three groups of 100 persons each. The correlation coefficients for total FAS scores were .97,.97 and .95. Test-retest reliability for Form B was studies in one instance to test reliability over time. This study yielded a coefficient of .82 for the total score. One study to assess equivalent form reliability yielded a coefficient of .86. No coefficients were reported for a second study of equivalent form reliability.
C.	What type of reliability coefficient was reported?	Split-half, test-retest, equivalent form.
D.	What was the composition of the group for whom reliability coefficients were computed?	Samplings from those in the standardization sample.
E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	Form A: FAS Verbal ranges from 2.5 to 4.1; Numerical ranges from 2.3 to 4.1; Total ranges from 3.4 to 5.1. Form B: FAS Verbal ranges from 3.9 to 4.0, Numerical ranges from 3.0 to 3.1; Total ranges from 4.8 to 5.4.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Fixed.

B.	What is the total time needed to score the tests?	About 10 minutes.
C.	Can the subtests be administered at different times?	Verbal and numerical tests can be administered separately.
D.	Will the test consume a reasonable portion of program/student time?	
E.	What qualifications are needed to administer the test?	Ability to operate a tape recorder. For scoring, the examiner should study the manual for accurate interpretation of scores.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Generally, yes. Directions are pre-taped. Assistance for students who do not follow directions is difficult.
H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	No.
J.	What is the range for the test series?	Designed to measure practical verbal and numerical ability of persons who have not completed school.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	Initial publication in 1968.
M.	What were the special adaptations, if any, for adults?	Especially designed for adults.
N.	What is the background of the authors and publishers?	Unknown.
O.	Is the format attractive and easy to follow?	Yes.

P.	Is the print size appropriate?	Type is of acceptable size for adult reading. Parts of the test are printed in all capital letters.
Q.	What are the components to be initially purchased?	Manual, booklets, scoring templates, and tapes for each form.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets must be repurchased.

Iowa Tests of Educational Development, Forms X-4 and Y-4

General Information:

1.	Title and level	Iowa Tests of Educational Development, Forms X-4 and Y-4
2.	Author	E.F. Lindquist and Leonard S. Feldt
3.	Publisher	Science Research Associates, 259 E. Erie Street, Chicago, Ill. 60611
4.	Date of publication	1963
5.	Cost	Full battery test booklets - $26.70/25; Individual test booklets - $4.05/25; IBM Answer Sheets - $36.90/500; (Examiner's Manuals and an answer key are included with answer sheets.)
6.	Time for administration	About 1 hour for each test
7.	Number of forms of the test	Two - X-4 and Y-4
8.	Type of test	Achievement
9.	Skill range	Grades 9-12
10.	Sections, parts, subtests	Test 1: Understanding of Basic Social Concepts Test 2: General Background in the Natural Sciences *Test 3: Correctness and Appropriateness of Expression *Test 4: Ability to do Quantitative Thinking *Test 5: Ability to Interpret Reading Materials in the Social Studies *Test 6: Ability to Interpret Reading Materials in the Natural Sciences *Test 7: Ability to Interpret Literary Materials Test 8: General Vocabulary Test 9: Use of Sources of Information

---

*Found to be of particular value for use in GED Programs due to their similarity to the five GED tests.

I. Validity:

A. What does the test measure (content)?

Test 1: Knowledge of facts involving historical developments before and after World War II; geography; economics; sociology; federal, state and local government: organization and powers; definition and understanding of social studies terms; and other topics.

Test 2: In the areas of biology, physics, chemistry, and earth sciences: ability to recall facts and definitions; to interpret concepts, laws, principals, facts, and conclusions; to recognize cause-and-effect relations, and draw conclusions; and other science-related skills.

Test 3: Ability to choose the most appropriate manner of expression, including colloquialisms and words often confused; grammar in context, including verb forms, pronoun forms, sentence structure, and word and sentence order; and writing skills including capitalization, paragraphing, conventions in letter writing, and especially punctuation. (Format is similar to GED Test format.)

Test 4: Ability to read and solve problems involving operations with fractions, percent, volume, area, units of measurement, ratio and proportion, averages, variability, angular relationships, evaluation of formulas, roots and powers, operations with signed numbers, cost and profit, interpretation of verbal statements, rate, linear interpolation, symbolic representation, solution of equations, verbalization of decimals, place value, number series, approximate computation and computational shortcuts, graphic representation, table reading, and graph reading. (No particular emphasis on any one of the above.)

Test 5: In the area of social studies: ability to comprehend what is stated in a selection (to locate and understand important facts, to restate ideas in new words and recognize valid examples, and to recognize the central thought and define the problem under discussion); ability to interpret what is implied in a selection (to evaluate the importance of ideas, to recognize simple relationships, to draw valid inferences, to extrapolate presented ideas to new but related situations, to deduce immediate conclusions and consequences); ability to analyze and evaluate a selection critically (to recognize generalized relationships between principal ideas; to demonstrate awareness of the writer's motives, approach, biases, argumentative techniques, and point of view; to judge the relevance of facts for the author's conclusions; to derive principal conclusions and generalizations from the selection, etc.)

Test 6: In the area of natural sciences: ability to comprehend the content of the selection (to restate ideas in new words and illustrate general principles, to recognize relationships, assess the importance of facts, summarize ideas, and draw valid inferences); ability to draw immediate conclusions from experimental results; ability to analyze and interpret an experiment as a whole (to define the problem of an experiment and to differentiate hypotheses, assumptions, date, and conclusions; to judge the relevance of facts for conclusions; to use background knowledge to explain aspects of an experiment; to derive principal conclusions and generalizations not directly stated, etc.)

Test 7: In reading literary materials: ability to comprehend and interpret the content of selections (to restate ideas in new words; to interpret figures of speech; to recognize simple relationships, draw inferences, and grasp the main thought of a selection; to understand the significance of specfic words or phrases in the light of literary context); ability to analyze and appreciate the elements of literary works (to understand the relationships between ideas and the function of individual ideas in relation to the total context; to grasp the structure of a selection and appreciate characterization; to evaluate an author's purpose, point of view, and attitudes; to recognize the style and literary techniques used in the development of a passage and appreciate its mood, tone, and emotion).

Test 8: Ability to recognize the meaning of words commonly encountered in reading by choosing one of four words that means the same as a given word in the context of a phrase.

Test 9: Ability to utilize appropriate sources of information, including the dictionary, thesaurus, library card catalog, encyclopedia, yearbook, almanac, compedia, periodicals, references pertaining to literary works, social practices and conventions, parts of a book, atlases, maps, globes, Readers' Guide to Periodical Literature and governmental sources of information.

Refer to Interpretive Supplement (Catalog #7-767) and How to Use the Test Results: A Manual for Teachers and Counselors (Catalog #7-758) for more detailed information on the content of each test.

B.	What type of validity does it exhibit?	Content, predictive, concurrent and construct.
	1. From what sources were the test items drawn?	Unknown.
	2. With what other tests or instruments was this test correlated?	a) Correlations of 9-Grade ITED. Composite Scores with ten other measure of high school achievement, including Graduation vs. Nongraduation, Rank in Graduating Class, Social Studies Average, Science Average, English Average, Mathematics Average, Grade 9 Grade-Point Average, and Cumulative Grade-Point Average through Grade 12. b) Correlations of ITED Composite Scores at various high school grade levels with College Freshman grades at nine institutions. c) Correlations of ITED Composite Scores with high school grades at various levels. Tests

	included are: ACE Psychological Examination, California Test of Mental Maturity, Henmon-Nelson Tests of Mental Ability, Otis Quick-Scoring Mental Ability Tests, Tests of Educational Ability. (Total Scores), and Primary Mental Abilities (Total Score).
What is the correlation coefficient?	a) The ten coefficients range from .45 to .73, with a median of .63. b) The seventeen coefficients range from .40 to .77, with a median of .61. c) The seven coefficients range from .50 to .72, with a median of .64. d) The twelve coefficients range from .572 to .848, with a median of .781.
What does the criterion test measure?	Tests mentioned in d) above are generally designed to measure aspects of intelligence rather than achievement.
3. What is the rationale for the test?	Designed to measure a number of highly general skills believed to be of lasting importance in adult life. They were not devised to serve as course examinations in the various areas of secondary school curriculum.
C. What skills does your program measure?	To be answered by program.
D. Examine the test itself.	
1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2. Are the test items well constructed?	To be answered by program.
3. Are they free from ambiguity?	To be answered by program.

E. What have reviewers, critics and users of the test said about the test?	Reviews from Buro's 6th Mental Measurements Yearbook: Ellis Batten Page states that ITED is one of the most widely-used measures of achievement ever designed. It is on some counts one of a number of nearly perfect instruments. Alexander G. Wesman criticizes the lack of information on ITED's validity and reliability, and the unsubstantiated claims for use of test results. However, he supports the publisher's claim that ITED provides a comprehensive and dependable description of the general educational development of the high school student, at least with respect to all broad aspects of educational development that are readily measurable. He suggests, however, that ITED is too comprehensive a test for yearly use, and that administration of the battery once in a student's career is sufficient. Additional comments: This test was designed and standardized for high school students. It has been found useful as a predictor of an adult student's ability to pass the GED test. Tests 3-7 of ITED are similar in format and content to the five GED tests. Experience has shown that a standard score of 13 on Tests 3-7 is a good predictor of a score of 45 on each of the corresponding GED Tests. (Follow the time limits for the full-length version in the ITED manual.)

II. Standardization Sample:

A. What was the composition of the persons on whom the test was standardized?	Public school students in grades 9-12. Participants were drawn from all geographic areas of the U.S. A stratified sample was chosen in accord with geographic region and size of secondary school enrollment.
B. How large was the sample?	51,098.

C.	From what geographic regions was the sample drawn?	New England, Mideast, Great Lakes, Plains, Southeast, Southwest, Rocky Mountains, Far West. Forty-nine states included.
D.	Are local norms available?	Local norms are provided to principals whose school has tests scored by SRA.
E.	In what terms are the norms reported?	Raw scores, standard scores, percentile ranks. Standard score scale ranges from 1-35 for use in assessing change.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Several are reported.
B.	What is the coefficient?	Ranges for subtests from .82 to .96. Median .90. For full battery .98 or .99.

	Range	Median
Test 1	.84-.93	.89
2	.85-.93	.90
3	.86-.93	.90
4	.82-.91	.88
5	.82-.92	.89
6	.85-.93	.91
7	.82-.92	.90
8	.92-.96	.94
9	.83-.92	.88
Full Battery		.98
Tests 1-8	.98-.99	

C.	What type of reliability was reported?	Split-half.
D.	What was the composition of the group for whom reliability coefficients were computed?	For the full battery: Answer sheets selected from those processed in the 1960 Iowa State High School Testing Program. For the partial battery: Selected from midwestern schools in 1960-62. Both included students from Grades 9-12. In all over 10,000 answer sheets were analyzed.

E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	Ranges from .99-.62 for the complete battery; from 1.21-2.71 for individual tests (includes full version and classroom version). Average for Tests 1-8 is 1.71 standard score units.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Fixed. The battery may be given either as a full length version requiring approximately one hour per test, or as a somewhat shorter class period version requiring about 40 minutes per test.
B.	What is the total time needed to score the test?	About 10 minutes. Machine-scoring is also available.
C.	Can the subtests be administered at different times.	Yes.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Ability to follow directions.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	Yes. However, this is done primarily in terms of high school students. See How to Use the Test Results: A Manual for Teachers and Counselors (Catalog # 7-758) and Manual for the School Administrator (Catalog # 7-1127).
I.	Are separate answer sheets available?	Yes.

J.	What is the range for the test series?	There is only one test level, spanning Grades 9-12.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	A revision of Form X-4, Form X-5, was published in 1971. Because Tests 5,6 and 7 have been combined into one test in the new edition, the test does not resemble the GED (High School Equivalency) Test to the extent that Form X-4 does.
M.	What were the special adaptations, if any, for adults?	None.
N.	What is the background of the authors and publishers?	Science Research Associates is a reputable test publisher which has produced a number of widely-used tests and other educational materials. Information on the authors is not readily available.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets, and answer sheets. Examiner's Manuals and Answer Keys accompany the answer sheets. Other supporting materials may be purchased.
R.	What parts of the test are reusable and which must be repurchased?	Answer sheets must be repurchased.

Metropolitan Achievement Tests, 1958 Edition

General Information:

1.	Title and level	Metropolitan Achievement Tests, Elementary, Intermediate, Advanced and High School Batteries, 1958 Edition
2.	Authors	Elementary, Intermediate and Advanced Batteries: Walter N. Durost, Harold H. Bixler, Gertrude H. Hildreth, Kenneth W. Lund and J. Wayne Wrightstone; High School Battery: Walter N. Durost (General Editor), William H. Evans, James D. Leake, Howard A. Bowman, Clark Cosgrove, John G. Read
3.	Publisher	Harcourt Brace Jovanovich, Inc., Test Department, 757 3rd Ave. New York, N.Y. 10017
4.	Date of publication	Elementary, Intermediate, Advanced: Form A - 1958; B - 1959; C - 1961; D - 1962; (Int. & Adv. only)
5.	Cost	Complete Battery Costs: Elementary: Consumable test booklets $11.00/35 Intermediate: Comsumable test booklets $14.00/35 Reusable test booklets $14.00/35 Answer sheets $8.40/35 sets Advanced: Consumable test booklets $14.50/35 Reusable test booklets $14.80/35 Answer sheets $8.40/35 sets High School: Reusable test booklets $13.90/35 Answer sheets $8.90/35 sets See catalog for prices of other components needed for the test. Partial battery booklets, as well as separate booklets for certain sub-tests, such as reading and math, are also available

6. Administration time

	Partial	Complete
Elementary	N.A.	2 Hrs., 27 Min.
Intermediate	3 Hrs., 17 Min.	3 Hrs., 57 Min.

	Partial	Complete
Advanced	3 Hrs., 27 Min.	4 Hrs., 7 Min.
High School	N.A.	5 Hrs., 15 Min.

Note: Additional time must be allowed for distributing and collecting materials and for giving directions

7.	Number of forms	Elementary - 3; Intermediate - 4; Advanced - 4; High School - 1
8.	Type of test	Achievement batteries
9.	Skill range	Grades 3-13: Elementary - Grades 3-4 Intermediate - Grades 5-6 Advanced - Grades 7-9 High School - Grades 9-13

10. Sections, parts, subtests

Elementary:
- Word Knowledge
- Word Discrimination
- Reading
- Spelling
- Language
  - Usage
  - Punctuation
  - Capitalization
- Arithmetic Computation
- Arithmetic Problem Solving & Concepts

Intermediate & Advanced Partial Battery:
- Word Knowledge
- Reading
- Spelling
- Language
  - Usage
  - Parts of Speech
  - Punctuation
  - Capitalization
- Language Study Skills
- Arithmetic Computation
- Arithmetic Problem Solving & Concepts
- Social Studies Study Skills

Complete Battery also includes Social Studies Information and Science

High School:
- Reading
- Spelling
- Language Arts
- Language Study Skills

Social Studies Study Skills
Social Studies Vocabulary
Social Studies Information
Mathematical Computation & Concepts
Mathematical Analysis & Problem Solving
Scientific Concepts & Understandings
Science Information

I. Validity:

A. What does the test measure?

Word Knowledge: ability to select from four or five choices the word which best defines a given word.

Word Discrimination: ability to select a word needed to complete a sentence from among several words of similar sound elements.

Reading: ability to correctly answer multiple choice questions about short paragraphs or longer selections.

Spelling: in the elementary battery, the ability to correctly spell words dictated orally; Intermediate, Advanced, High School: ability to recognize whether or not a word used in context is correctly or incorrectly spelled; if incorrect, the ability to write the correct spelling.

Language Usage: words, punctuation, capitalization are used within a sentence context. Testees must identify whether or not underlined items are used correctly or incorrectly. If incorrect, the testee must write the correct form or identify how to correct the error. A section on the advanced tests also requires the ability to identify the part of speech of an underlined word or phrase in a sentence.

Language Study Skills: ability to decide the correct dictionary definition for words used in context; ability to exhibit knowledge of other dictionary skills - syllabication, use of guide words,

		accent, letter sounds, alphabetical order; ability to identify the best source of information for given topics.
		Social Studies Study Skills: ability to categorize social studies words according to a few broad topics.
		Social Studies Information: ability to select the best of four possible answers to short questions about social studies.
		Scientific Concepts & Understanding: ability to categorize science terms under a broader concept term. Also, ability to select words which will complete short paragraphs relating to science.
		Science Information: ability to select the best of four possible answers to short questions about science.
		Arithmetic Computation: ability to complete arithmetic problems.
		Mathematical Computation & Concepts: ability to select the correct answer for a math problem asked in sentence format.
		Arithmetic Problem Solving & Concepts: ability to read and correctly select one of four possible answers to questions related to arithmetic; ability to read, decipher and compute word problems.
		Mathematical Analysis & Problem Solving: ability to read, decipher and compute word problems.
B.	What type of validity does it exhibit?	Content validity.
	1. From what sources were the test items drawn?	Inclusion of items was based on a detailed outline of material from a wide variety of sources that cited important goals and outcomes subject by subject, and grade by grade. The outline is based on an analysis of textbooks, courses of study, and expert formulations of the goals of instruction in the elementary and secondary education.
	2. With what other tests or instruments was this test correlated?	Information not available.

	What is the correlation coefficient?	Information not available.
	What does the criterion test measure?	Information not available.
	3. What is the rationale for the test?	The tests attempt to measure those outcomes of instruction which, according to authoritative judgment and consensus of current practice, are the important goals of elementary and secondary education.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Reviews from Buro's 6th & 7th Mental Measurements Yearbooks by Paul L. Dressel, Henry S. Dyer, Warren G. Findley, H. Alan Robinson, C.F. Anderhalter, E.W. Hamilton, Elizabeth Hagen and Frank B. Womer, indicate: -The tests are lacking in the number of items that require critical thinking and problem solving. Most of the questions on the Arithmetic Problem Solving and Concepts tests are very routine verbal problems; not enough emphasis has been placed on concepts, generalizations, and understanding. -The tests have not been sufficiently up-dated to reflect current elementary and secondary curriculum. -Neither in content level nor in the range of cognitive objectives are the tests adequate to provide direction to teachers in the improvement of instruction and the curriculum.

-Norms are not provided for the total scores for each of the large blocks of the curriculum, language arts and mathematics.

-The Manual for Interpreting approaches in quality and scope is a textbook in educational measurement, with particular application to the Metropolitan Achievement Tests. It is exceptionally well done.

-Although all tests at all levels provide ample scales for reflecting superior achievement, measurement of poor learners is generally inadequate. For most subtests, the minimum grade level for which norms are offered is not much lower than the score that could be obtained by random guessing.

-The Metropolitan Achievement Test is to be applauded for scope, both vertical and horizontal.

-The use of DK (Don't Know) response option in the spelling and several of the language usage sections of the tests is of questionable value, and can be the cause of distorted scores, depending upon each tester's instructions in this regard.

-The test booklets are attractive and item positioning, type size and coloring add to the readability.

Additional Comments:

-Items are clearly school and textbook oriented. Care must be exercised in using these tests with adults, particularly the lower grade levels, since the language is more child oriented.

-The tests were designed and normed for children.

-The fact that test content is outdated may make the test more appropriate for use with adults, since it is more likely to reflect their formal education.

II. Standardization Sample:

A. What was the composition of the persons on whom the test was standardized?	For Elementary, Intermediate and Advanced levels, factors of school size, geographic location, type of community, student intelligence,

		and type of school system were considered in selecting the random sample. Students in 10th, 11th, and 12th Grades, randomly selected from communities having populations between 10,000 and 100,000, were included for the High School sample.
B.	How large was the sample?	The Primary through Advanced test sample encompassed over 500,000 students in 225 school systems. The High School battery included more than 31,000 students in 29 school systems.
C.	From what geographic regions was the sample drawn?	For the Elementary through Advanced tests, the sample was drawn from 49 states representing four regions: 1) New England and Middle Atlantic; 2) North Central; 3) Southern; and 4) Pacific Coast. The High School test sample was drawn from 19 states representing the above four regions.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported?	At each battery level, the raw scores on all subtests are converted to a set of normalized standard scores. Percentiles and stanines can be derived from standard scores.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is the coefficient?	Ranges for all sub-tests

Ranges for all sub-tests	
Elementary	.61-.95
Intermediate	.64-.96
Advanced	.53-.95
High School	.72-.99

C.	What type of reliability was reported?	For Elementary through Advanced: Split-half. For High School: Split-half; alternate form.
D.	What was the composition of the group for whom reliability coefficients were computed?	Elementary: Four groups of 100 each from grade 4.1 in a single school. Intermediate: Four groups of 100 each from grade 6.1 in a single school. Advanced: Four groups of 100 each from grade 8.1 in a single school. High School: Split-half reliability. Two groups of 100 each from grades 10 and 11 from two schools. Alternate form reliability: Two estimates based on a random sample (N=393-585 for subtests) from two schools.
E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	Elementary: Median of raw scores for subtests ranges from 1.7 to 3.5. Intermediate: Median of raw scores for subtests ranges from 1.3 to 3.0. Advanced: Median of raw scores for subtests ranges from 1.3 to 3.5. High School: Form A: 1.4 to 3.5 raw score points of subtests.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Fixed.
B.	What is the total time needed to score the test?	Ten minutes if answer sheets are used, thirty minutes if not. Machine scoring available.
C.	Can the subtests be administered at different times?	Yes, but a procedure is suggested for administering subtests in a series of sessions.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.

E.	What qualifications are needed to administer the test?	No specific qualifications are needed. Explicit instructions are included for all tests and subtests.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Yes. Directions are well written.
H.	Does the manual contain guides for using and interpreting the scores?	Yes. Each test has its own guides.
I.	Are separate answer sheets available?	Separate answer sheets are available for the Intermediate, Advanced and High School batteries.
J.	What is the range for the test series?	Latter half of grade 1 through 12.
K.	What is the origin of the test?	The first edition was published in 1932.
L.	How recently has the test been revised?	There is a new 1970 edition available.
M.	What were the special adaptations, if any, for adults?	None.
N.	What is the background of the authors and publishers?	Walter N. Durost, General Editor: Director, Division of Research and Test Service, World Book Co.; Associate Professor of Education, Boston University; Director of Educational Services, Pinellas County, Florida; Special Consultant, Test Department, Harcourt, Brace & World. Publisher has wide and varied experience in books and tests for children through adults.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.

Q.	What are the components to be initially purchased?	Test booklets, scoring keys, and manuals for administering. For machine scoring, separate answer sheets will have to be purchased for Intermediate, Advanced, and High School batteries.
R.	What parts of the test are reusably and which must be repurchased?	Consumable booklets are available for Elementary, Intermediate, and Advanced batteries. Reusable booklets are available for Intermediate, Advanced, and High School batteries. Answer sheets, if used, must be repurchased.

Tests of Adult Basic Education, Level E

General Information:

1.	Title and level	Tests of Adult Basic Education (TABE), Level E
2.	Author	Ernest W. Teigs and Willis W. Clark
3.	Publisher	California Test Bureau, A Div. of McGraw-Hill Book Co., Del Monte Research Park, Monterey, Calif. 93940
4.	Date of publication	1967 (Adapted from 1957 Edition of California Achievement Tests, Upper Primary Level, Forms W and X, 1963 norms)
5.	Cost	Test Booklets: (Scoring Key and Administrator's Manual included) $6.00/25; Answer Sheets: IBM 1230: $1.50/25 (for each subject) or SCOREZE: $2.50/25 (for each subject); Profile Sheets: $1.00/25
6.	Time for administration	Time limits of subtests total 1 hour, 34 minutes. Additional time must be allowed for giving directions, distributing materials, etc.
7.	Number of forms of the test	Two - Form 1 and Form 2
8.	Type of test	Battery of achievement tests
9.	Skill range	Grade levels 2-3
10.	Sections, parts, subtests	Reading: Vocabulary (2 sections) Comprehension (3 sections) Arithmetic Reasoning (4 sections) Fundamentals (4 sections)

I. Validity:

Validity data is based on the California Achievement Tests, Upper Primary Level, unless otherwise indicated.

A.	What does the test measure (content)?	Reading Vocabulary: 1) word recognition (choosing the word pronounced

by the tester from 3 words with differing initial, middle, or final sounds), and 2) word meaning (matching words with their opposites).

Reading Comprehension: Following simple directions, following directions requiring choice, knowing and using parts of a book, reading a bar graph, alphabetizing, understanding main ideas and sequences, understanding directly stated facts, and drawing conclusions. Mostly multiple choice.

Arithmetic Reasoning: Writing and recognizing number values (including Roman numerals and money), telling time to the half hour, knowing the value of currency, knowing the meaning of operational signs and abbreviations of terms used in measurement, and working one and two-step word problems involving small whole numbers and money.

Arithmetic Fundamentals: Primarily on basic number facts of the four operations, but also includes the use of zero; and of borrowing and carrying in operations with whole numbers.

B.	What type of validity does it exhibit?	Criterion.
	1. From what sources were the test items drawn?	From a consensus of professional judgment about what constitutes the basic skills in reading and arithmetic. Details are not given.
	2. With what other tests or instruments was this test correlated?	California Short-Form Test of Mental Maturity, 1963 Revision. (CTMM-SF, 1H, Grade 3).
	What is the correlation coefficient?	Total Reading: .81; Total Arithmetic: .72; Total Battery: .81. Coefficients are also given for subtests ranging from .58 to .79. None are given for the sections. The Technical Report also shows correlation with parts of the Iowa Tests and Basic Skills and Metropolitan Achievement Tests.

	What does the criterion test measure?	The California Short-Form Test of Mental Maturity is divided into Language and Non-Language sections each yielding an IQ and a mental age. Standard scores, percentile ranks, and stanines are also provided in the norms tables for four factors: Logical Reasoning, Numerical Reasoning, Verbal Concepts, and Memory.
3.	What is the rationale for the test?	TABE is designed to measure basic skills, not content of the various school subjects. An analysis of learning difficulties provides information which can enable the student to progress toward greater proficiency in using the basic number and language skills required of him daily in our society. Some changes were made in revising the test in order to make it more appealing to adults.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Buro's 7th Mental Measurements Yearbook includes reviews by A.N. Hieronymous and S. Alan Cohen. Both reviewers express concern over the fact that the TABE manuals contain no technical information on TABE or CAT, that norms and other technical information were not revised after revision of the test, and that no adults are included in the norming population.

The question whether the content is any more appropriate for adults than are most other batteries developed for use in elementary schools. Other comments on TABE include: 1) The two alternate forms for each level should be regarded as paraphrased rather than equivalent (Hieronymous). 2) The Analysis of Learning Difficulties encourages the use to interpret raw scores derived from small groups of items. These scores are completely meaningless (Hieronymous). 3) The Practice Exercises and Locator Test is a most useful provision that might reduce the effects of diverse cultural background (Cohen).

II.	Standardization Sample:	Standardization data is based on the California Achievement Tests, Upper Primary Level, unless otherwise indicated. No data is available for TABE-E.
A.	What was the composition of the persons on whom the test was standardized?	Norms are based on children, reportedly chosen according to a research plan designed to control bias from any one section of the country or any one type of educational program or school system. Details of the research plan are not included in the Technical Report.
B.	How large was the sample?	Between 1,000 and 1,500 from each grade (Grades 2 and 3), i.e., between 2,000 and 3,000 children. Exact numbers are not reported.
C.	From what geographic region was the sample drawn?	Reportedly from a variety of regions of the United States. Details are not included in the Technical Report.
D.	Are local norms available?	
E.	In what terms are the norms reported (grade level, age, etc.)?	For TABE, in raw scores converted to grade equivalents.
F.	Are students similar to those in your program represented in the sample?	

III. Reliability: Reliability information is based on California Achievement Tests, Upper Primary Level, Form W, which corresponds to Form 1 of TABE E. No reliability data is available for other forms of the test, or for TABE E.

A. Is a reliability coefficient reported? Yes.

B. What is this coefficient?

Test	Internal Consistency (Split-half)	Test-Retest
Total Reading	.94	.73
Total Arithmetic	.96	.44
Total Battery	.98	.61

Split-half coefficients are given for subtests, ranging from .88 to .95. Test-retest coefficients given for subtests range from .38 to .71. No coefficients are given for the sections.

C. What type of reliability coefficient was reported?

1) Internal consistency (split-half) coefficients.
2) Test-retest coefficients.

(See supplementary page on TABE, p.118, for inter-level correlations.)

D. What was the composition of the group for whom reliability coefficients were computed?

1) Split-half coefficients are based on 131 primary students in a single system, Grade 2.7.
2) Test-retest coefficients are based on two administrations of the same test to 90 students in a single system, Grades 2.8 and 3.8.

E. Is a standard error of measurement reported? Yes.

F. What is it? Total Reading: 0.2; Total Arithmetic: 0.2; Total Battery: 0.1; SE's are given for subtests, ranging from 0.1 to 0.2. None are given for the sections.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Fixed.
B.	What is the total time needed to score the test?	A minimum of 15 minutes. Considerably more time is required for transferral of information to profile sheet and use of the Analysis of Learning Difficulties.
C.	Can the subtests be administered at different times?	Yes, a break may be taken after any timed subsection. It is recommended that the Reading and Arithmetic sections be administered in 2 different sittings, or that at least a break be given between the sections.
D.	Will the test consume a reasonable portion of the program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	No special qualifications, except ability to follow directions exactly as given in the manual.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Students who answer less than 5 items correctly on the Locator Test will have difficulty because they cannot read. Such students should not be given the test. Most students on this level have difficulty answering multiple-choice questions. Extreme care should be taken in giving directions and giving individual assistance in the mechanics of the test to students who need help.
H.	Does the manual contain guides for using and interpreting the scores?	Very little. The Analysis of Learning Difficulties, found on the reverse side of the student profile sheet, is helpful in determining specific difficulties of the student.
I.	Are separate answer sheets available?	No.

J.	What is the range for the test series?	Grade levels 2-9.
K.	What is the origin of the test?	Consensus of professional judgment about what constitutes the basic skills in reading and mathematic.
L.	How recently has the test been revised?	Tests of Adult Basic Education, published in 1967 through minor revisions of the *California Achievement Test*, has not been revised since that time. *California Achievement Tests* were revised in 1970.
M.	What were the special adaptations, if any, for adults?	Some of the juvenile vocabulary was elminated, the Language Test omitted and the title and cover of the test was changed.
N.	What is the background of the authors and publishers?	California Test Bureau is a reputable test publisher, having produced a number of widely-used tests. The authors have designed a number of the CTB's tests.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets, profile sheets (if desired), answer keys, Practice Exercises and Locator Test (if desired), and Examiner's Manuals.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets, profile sheets, and Practice Exercises and Locator Tests are consumable. Answer keys and Examiner's Manuals are reusable.

Tests of Adult Basic Education, Level M

General Information:

1.	Title and level	Tests of Adult Basic Education (TABE), Level M
2.	Author	Ernest W. Tiegs and Willis W. Clark
3.	Publisher	California Test Bureau, A Div. of McGraw-Hill Book Co., Del Monte Research Park, Monterey, Calif. 93940
4.	Date of publication	1967. (Adapted from 1957 Edition of *California Achievement Tests*, Elementary Level, Forms W and X, 1963 norms)
5.	Cost	Test Booklets: $7.40/25 (Administrator's Manual included); Answer Sheets: IBM 1230: $1.50/25 (for each subject) or SCOREZE: $2.50/25 (for each subject); Profile Sheets: $1.00/25; Scoring Stencils: $2.25/set
6.	Time for administration	Time limits of subtests total 2 hours, 38 minutes. Additional time must be allowed for giving directions, distributing materials, etc.
7.	Number of forms of the test	Two - Form 1 and Form 2
8.	Type of test	Battery of Achievement Tests
9.	Skill range	Grade levels 4-6
10.	Sections, parts, subtests	

Reading	Arithmetic	Language
Vocabulary (4 sections) Comprehension (3 sections)	Reasoning (3 sections) Fundamentals (4 sections)	Mechanics of English (3 sections) Spelling (1 section)

I. Validity: Validity data is based on the *California Achievement Tests*, Elementary Level, unless otherwise indicated.

A.	What does the test measure (content)?	Reading Vocabulary: Choosing one of 4 words that means the opposite of a given word. Involves math, science, social science, and general vocabulary. Reading Comprehension: Following directions, reference skills (including knowing the parts of a book, reading graphs and maps, alphabetizing, using an index), getting the main idea, sequence of events, drawing conclusions, and finding facts. Nearly all multiple-choice. Arithmetic Reasoning: Changing words and Roman numerals to Arabic numerals, choosing the smallest of 4 numbers (involving whole numbers, fractions, decimals, and percents), knowing symbols and abbreviations involved in operations and measures, one and two-step word problems primarily involving whole numbers and measures. All multiple-choice. Arithmetic Fundamentals: Computations primarily involving whole numbers and fractions. All multiple-choice. Mechanics of English: Capitalization (indicating words in a sentence that should be capitalized - proper nouns and first words of sentences and quotations); punctuation (telling whether a period, comma, question mark, quotation mark, or no punctuation is needed in specified places in a letter and in a story); usage (choosing one of two words that should be used in a sentence - primarily correct forms of pronouns and verbs - and recognizing complete sentences); Spelling (choosing incorrectly spelled words). All multiple-choice.
B.	What type of validity does it exhibit?	Criterion.
	1. From what sources were the test items drawn?	From a consensus of professional judgment about what constitutes the basic skills in reading, language, and arithmetic. Details are not given.

	2. With what other tests or instruments was this test correlated?	California Short-Form Test of Mental Maturity, 1963 Revision, (CTMM-SF-2, Grade 4).
	What is the correlation coefficient?	Total Reading: .80; Total Arithmetic: .63; Total Language: .74; Total Battery: .81. Coefficients are also given for subtests, ranging from .43 to .78. None are given for the sections. The Technical Report also shows correlation with parts of the Iowa Tests of Basic Skills and Metropolitan Achievement Tests.
	What does the criterion test measure?	The California Short-Form Test of Mental Maturity is divided into Language and Non-Language sections each yielding an IQ and a mental age. Standard scores, percentile ranks, and stanines are also provided in the norms tables for four factors: Logical Reasoning, Numerical Reasoning, Verbal Concepts, and Memory.
	3. What is the rationale for the test?	TABE is designed to measure basic skills not content of the various school subjects. An analysis of learning difficulties provides information which can enable the student to progress toward greater proficiency in using the basic number and language skills required of him daily in our society. Some minor changes were made in revising the test in order to make it more appealing to adults.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.

E.	What have reviewers, critics and users of the test said about the test?	See summary of Buro's reviews for TABE E, p. 99. Additional comment: The advantage of this test over its competitors is the inclusion of a language test (Hieronymous).
II.	Standardization Sample:	Standardization data is based on the California Achievement Tests, Elementary Level, unless otherwise indicated.
A.	What was the composition of the persons on whom the test was standardized?	Norms are based on children, reportedly chosen according to a research plan designed to control bias from any one section of the country or any one type of educational program or school system. Details of the research plan are not included in the Technical Report.
B.	How large was the sample?	Between 1,000 and 1,500 from each grade (Grades 4,5, and 6) i.e., between 3,000 and 4,500 children. Exact numbers are not reported.
C.	From what geographic regions was the sample drawn?	Reportedly from a variety of regions of the United States. Details are not included in the Technical Report.
D.	Are local norms available?	
E.	In what terms are the norms reported (grade level, age, etc.)?	For TABE, in raw scores converted to grade equivalents.
F.	Are students similar to those in your program represented in the sample?	
III.	Reliability:	Reliability information is based on California Achievement Tests, Elementary Level, Form W, which corresponds to Form 1 of TABE. No reliability data is available for other forms of the test.
A.	Is a reliability coefficient reported?	Yes.

B. What is this coefficient?

Test	Split-half	Test-Retest	
		4.8-5.8	5.8-6.8
Total Reading	.95	.86	.88
Total Arithmetic	.87	.84	.83
Total Language	.96	.84	.83
Total Battery	.97	.91	.92

Split-half coefficients are given for subtests, ranging from .78 to .95. Test-retest coefficients for subtests range from .70 to .84. No coefficients are given for the sections.

C.	What type of reliability coefficient was reported?	1. Internal consistency (split-half) coefficients. 2. Test-retest coefficients. (See supplementary page on TABE, page for inter-level correlation.)
D.	What was the composition of the group for whom reliability coefficients were computed?	1. Split-half coefficients are based on 384 elementary school students in a single system, Grade 5.1. 2. Test-retest coefficients are based on two administrations of the same test to two groups of students, each in a single system: 90 students, Grades 4.8 and 5.8, and 125 students, Grades 5.8 and 6.8.
E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	Total Reading: 0.3; Total Arithmetic: 0.3; Total Language: 0.2; Total Battery: 0.2; SE's are given for subtests, ranging from 0.3 to 0.5. None are given for the sections.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Fixed.
B.	What is the total time needed to score the test?	A minimum of 15 minutes per test if Scoreze or IBM Answer Sheets with transparent scoring stencils are used; longer if answers are written in the booklets. For

		transferral of information to the profile sheet and use of the Analysis of Learning Difficulties considerably more time is required. Machine scoring is available.
C.	Can the subtests be administered at different times?	Yes, a break may be taken after any timed subsection. It is recommended that the Reading, Arithmetic, and Language sections be administered in 3 different sittings, or that at least a break be given between the sections.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	No special qualifications, except ability to follow directions exactly as given in the manual. S. Alan Cohen states, "Group administered California tests have about the easiest administration instructions of all standardized tests of this kind." (See Buro's 7th Mental Measurements Yearbook.)
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Use of IBM Answer Sheets can be difficult for some adult students. Administration of Practice Exercises and Locator Test should eliminate most problems on this level, if directions are given properly.
H.	Does the manual contain guides for using and interpreting the scores?	Very little. The Analysis of Learning Difficulties found on the reverse side of the profile sheet is helpful in determining possible areas of difficulty for each student. No technical data is included in the TABE Manual.
I.	Are separate answer sheets available?	Scoreze and IBM 1230 Answer Sheets with transparent scoring stencils are available. Answers may also be written in the test booklets.
J.	What is the range for the test series?	Grade levels 2-9.

K.	What is the origin of the test?	Consensus of professional judgment about what constitutes the basic skills in reading, arithmetic, and language.
L.	How recently has the test been revised?	Tests of Adult Basic Education, published in 1967 through minor revision of the California Achievement Tests, has not been revised since that time. California Achievement Test was revised in 1970.
M.	What were the special adaptations, if any, for adults?	Some of the juvenile vocabulary was eliminated, and the title and cover of the test was changed.
N.	What is the background of the authors and publishers?	California Test Bureau is a reputable test publisher, having produced a number of widely-used tests. The authors have designed a number of CTB's tests.
O.	Is the format attractive and easy to follow?	Answer sheets are sometimes difficult for students to follow.
P.	Is the print size appropriate?	Type on IBM Answer Sheets is too small for some students.
Q.	What are the components to be initially purchased?	Test booklets, answer sheets (if desired), profile sheets (if desired), answer keys, Practice Exercises and Locator Test (if desired), and Examiner's Manuals.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets, may be reused if answer sheets are used. Answer keys and Teacher's Manuals are reusable. Other components are consumable.

Tests of Adult Basic Education, Level D

General Information:

1.	Title and level	Tests of Adult Basic Education (TABE), Level D
2.	Author	Ernest W. Tiegs and Willis W. Clark
3.	Publisher	California Test Bureau, A Div. of McGraw-Hill Book Company, Del Monte Research Park, Monterey, Calif. 93940
4.	Date of publication	1967. (Adapted from 1957 Edition of California Achievement Tests, Junior High Level, Forms W and X, 1963 norms)
5.	Cost	Test Booklets: $7.40/25 (Administrator's Manual included); Answer Sheets: IBM 1230: $1.50/25 (for each subject) or SCOREZE: $2.50/25 (for each subject); Profile Sheets: $1.00/25; Scoring Stencils: #2.25/set
6.	Time for administration	Time limits of subtests total 2 hrs., 56 min. Additional time must be allowed for giving directions, distributing materials, etc.
7.	Number of forms of the test	Two - Form 1 and Form 2
8.	Type of test	Battery of achievement tests
9.	Skill range	Grade levels 7-9
10.	Sections, parts, subtests	

Reading	Arithmetic	Language
Vocabulary (4 sections)	Reasoning (3 sections)	Mechanics of English (3 sections)
Comprehension (3 sections)	Fundamentals (4 sections)	Spelling (1 section)

I. Validity: Validity data is based on the California Achievement Tests, Junior High Level, unless otherwise indicated.

A.	What does the test measure (content)?	Reading Vocabulary: Choosing one of 4 words that means the opposite of a given word. Involves math, science, social science, and general vocabulary. Reading Comprehension: Following directions, reference skills (including knowing and using the parts of a book, reading graphs and maps, using an index and the table of contents, and using the library and encyclopedia), getting the main idea, sequence of events, drawing conclusions, and finding facts. All multiple-choice. Arithmetic Reasoning: Changing words and Roman numerals to Arabic numerals, choosing the smallest of 4 numbers (involving whole numbers, fractions, decimals, and percents), knowing symbols and abbreviations involved in operations and measures, knowing formulas, operations with positive and negative numbers, solving simple algebraic equations; and word problems, mostly two-step, primarily involving whole numbers, measures, and percentage. Arithmetic Fundamentals: Computations primarily involving whole numbers, fractions and decimals. Mechanics of English: Capitalization (indicating words in a sentence that should be capitalized--proper nouns and first words of sentences and quotations); punctuation (telling whether a comma, question mark, single or double quotation mark, apostrophe, or no punctuation is needed in specified places in a letter and in a story); usage (choosing one of two words that should be used in a sentence--primarily correct forms of pronouns and verbs--knowing parts of speech, and recognizing complete sentences); spelling (choosing incorrectly spelled words). All multiple-choice.
B.	What type of validity does it exhibit?	Criterion.

1.	From what sources were the test items drawn?	From a consensus of professional judgment about what constitutes the basic skills in reading, language and arithmetic. Details are not given.
2.	With what other tests or instruments was this test correlated?	*California Short-Form Test of Mental Maturity*, 1963 Revision. (CTMM-SF-3, Grade 9).
	What is the correlation coefficient?	Total Reading: .83; Total Arithmetic: .74; Total Language: .67; Total Battery: .83; Coefficients are also given for subtests, ranging from .46 to .82. None are given for the sections.
	What does the criterion test measure?	The *California Short-Form Test of Mental Maturity* is divided into Language and Non-language sections each yielding an IQ and a mental age. Standard scores, percentile ranks, and stanines are also provided in the norms tables for four factors: Logical Reasoning, Numerical Reasoning, Verbal Concepts, and Memory.
3.	What is the rationale for the test?	TABE is designed to measure basic skills, not content of the various school subjects. An analysis of learning difficulties provides information which can enable the student to progress toward greater proficiency in using the basic number and language skills required of him daily in our society. Some minor changes were made in revising the test in order to make it more appealing to adults.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.

2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	See summary of Buro's reviews for TABE E, page 99. Additional comment: The advantage of this test over its competitors is the inclusion of a language test (Hieronymous).
II.	Standardization Sample:	Standardization data is based on the *California Achievement Tests*, Junior High Level, unless otherwise indicated.
A.	What was the composition of the persons on whom the test was standardized?	Norms are based on children, reportedly chosen according to a research plan designed to control bias from any one section of the country or any one type of educational program or school system. Details of the research plan are not included in the *Technical Report*.
B.	How large was the sample?	Between 1000 and 1500 from each grade (Grades 7,8 and 9), i.e., between 3000 and 4500 children. Exact numbers are not reported.
C.	From what geographic regions was the sample drawn?	Reportedly from a variety of regions of the United States. Details are not included in the *Technical Report*.
D.	Are local norms available?	
E.	In what terms are the norms reported (grade level, age, etc.)?	For TABE, in raw scores converted to grade equivalents.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.
III.	Reliability:	Reliability information is based on *California Achievement Tests*, Junior High Level, Form W, which corresponds to Form 1 of TABE. No reliability data is available for other forms of the test.

A. Is a reliability coefficient reported? — Yes.

B. What is this coefficient?

Test	Split-half	Test-Retest	
		7.8-8.8	8.8-9.8
Total Reading	.95	.95	.92
Total Arithmetic	.94	.93	.90
Total Language	.93	.91	.91
Total Battery	.98	.97	.95

Split-half coefficients are given for subtests, ranging from .83 to .93. Test-retest coefficients for subtests range from .82 to .93. No coefficients are given for the sections.

C. What type of reliability coefficient was reported?

1. Internal consistency (split-half) coefficients.
2. Test-retest coefficients. (See supplementary page on TABE page for inter-level correlations.)

D. What was the composition of the group for whom reliability coefficients were computed?

1. Split-half coefficients are based on 200 junior high school students in a single system, Grade 8.1.
2. Test-retest coefficients are based on two administrations of the same test to two groups of students, each in a single system: 107 students, Grades 7.8 and 8.8, and 100 students, Grades 8.8 and 9.8.

E. Is a standard error of measurement reported? — Yes.

F. What is it? — Total Reading: 0.4; Total Arithmetic: 0.4; Total Language: 0.5; Total Battery: 0.2; SE's are given for subtests, ranging from 0.5 to 0.7. None are given for the sections.

IV. Practicality:

A. In administering the tests, are time limits fixed or flexible? — Fixed.

B.	What is the total time needed to score the test?	A minimum of 15 minutes per test if SCOREZE or IBM Answer Sheets with transparent scoring stencils are used; longer if answers are written in the booklets. For transferral of information to the profile sheet and use of the Analysis of Learning Difficulties considerably more time is required. Machine scoring is available.
C.	Can the subtests be administered at different times?	Yes, a break may be taken after any timed subsection. It is recommended that the Reading, Arithmetic and Language sections be administered in 3 different sittings, or that at least a break be given between the sections.
D.	Will the test consume a reasonable portion of the program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	No special qualifications, except ability to follow directions exactly as given in the manual. S. Alan Cohen states, "Group administered California tests have about the easiest administration instructions of all standardized tests of this kind." (See Buro's 7th Mental Measurements Yearbook.)
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Use of IBM Answer Sheets can be difficult for some adult students. Administration of Practice Exercises and Locator Test should eliminate all problems on this level, if directions are given properly.
H.	Does the manual contain guides for using and interpreting the scores?	Very little. The Analysis of Learning Difficulties found on the reverse side of the profile sheet is helpful in determining possible areas of difficulty for each student. No technical data is included in the TABE manual.
I.	Are separate answer sheets available?	SCOREZE and IBM 1230 answer sheets with transparent scoring stencils

		are available. Answers may also be written in the test booklets.
J.	What is the range for the test series?	Grade levels 2-9.
K.	What is the origin of the test?	Consensus of professional judgment about what constitutes the basic skills in reading, arithmetic, and language.
L.	How recently has the test been revised?	Test of Adult Basic Education, published in 1967 through minor revisions of the California Achievement Tests, has not been revised since that time. California Achievement Tests were revised in 1970.
M.	What were the special adaptations, if any, for adults?	A few insignificant changes were made in content and wording. The title and cover of the test was changed.
N.	What is the background of the authors and publishers?	California Test Bureau is a reputable test publisher, having produced a number of widely-used tests. The authors have designed a number of CTB's tests.
O.	Is the format attractive and easy to follow?	Answer sheets are sometimes difficult for some students to follow.
P.	Is the print size appropriate?	Print size on IBM answer sheets is too small for some students.
Q.	What are the components to be initially purchased?	Test booklets, answer sheets (if desired), profile sheets (if desired), answer keys, Practice Exercises and Locator Test (if desired), Examiner's Manuals.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets may be reused if answer sheets were used. Answer keys and Examiner's Manuals are reusable. Other components are consumable.

SUPPLEMENTARY PAGE FOR TABE

A. Practice Exercises and Locator Test

The Practice Exercises are designed to provide experiences with the mechanics of marking answers to objective test items on separate answer sheets, to develop some test-taking sophistication, and to minimize the effects of diverse backgrounds of experience in the use of objective tests and separate sheets.

The Locator Test is a short vocabulary test used to determine the appropriate level of TABE for each individual. A transparent hand-scoring stencil makes it possible to score the answer sheets for a large group of students in less than 15 minutes so that the examiner can readily determine the appropriate level to administer to a particular examinee. Cut-off scores for each level of the TABE series are provided in the section labeled "Interpretation of Scores."

Cost: Tests $2.50/25 (Administrator's manual included); Answer Sheets: $1.25/25; Scoring Stencils: 75¢ each.

B. Inter-level Correlations

Inter-level coefficients show the degree of equivalence of test scores with adjacent levels. They determine the feasibility of choosing a higher or lower test for retesting. One hundred, nineteen students of third grade were tested on the Upper Primary Level, and a year later tested on the Elementary Level. One hundred, three sixth grade students were tested on the Elementary Level, and a year later on the Junior High Level. Results are as follows:

	Up. Pri. (Level E) (Grade 3.8) Elem. (Level M) (Grade 4.8)	Elem. (Level M) (Grade 6.8) Jr. H. (Level D) (Grade 7.8)
Total Reading	.73	.89
Total Arithmetic	.58	.87
Total Language	.66	.82
Total Battery	.72	.93

Coefficients for subtests on Level E/Level M range from .42 to .73; on Level M/Level D range from .75 to .93. This data shows that it is not statistically acceptable to compare scores on the Level E test to scores on the Level M test, while it is fairly acceptable to compare Level M to Level D.

Wide Range Achievement Test

General Information:

1.	Title and level	The Wide Range Achievement Test
2.	Author	J. F. Jastak and S. R. Jastak
3.	Publisher	Guidance Associates 1526 Gilpin Avenue Wilmington, Delaware
4.	Date of Publication	1965, revised
5.	Cost	Test forms (pkg. 50): $4.25 Manual: $3.00 Specimen Set: $3.15
6.	Time for Administration	Between 20 and 30 minutes
7.	Number of forms of the test	One
8.	Type of test	Achievement
9.	Skill Range	Pre-school through college
10.	Sections, Parts, Subtests	Three: Spelling, Arithmetic, and Reading

I. Validity:

A.	What does the test measure (content)?	Reading: The ability to recognize and name letters, and pronounce words. Spelling: The ability to copy marks resembling letters, writing the name, and writing single words to dictation. Arithmetic: The ability to count and read number symbols, solve oral problems, and perform written computations.
B.	What type of validity does it exhibit?	Criterion validity.
1.	From what sources were the test items drawn?	Not stated.

	2. With what other tests or instruments was this test correlated?	Numerous comparisons have been made with intelligence tests such as the Wechaler Intelligence Scales; achievement tests such as the California Mental Maturity; and reading tests such as the Wood-Saugren and Stanford.
	3. What is the rationale for the test?	Designed as a measure for the basic school subjects of reading, spelling and arithmetic.
C.	What skills does your program measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers and critics and users of the test said about the test?	Two reviews of WRAT, by Jack C. Merwin and Robert L. Thorndike, are included in Buros' Seventh Mental Measurements Yearbook. Both reviewers question its use as an achievement test because of its limited content. It provides only a rough indication of three limited components of educational achievement. The individual user of the test must determine whether these components reflect "achievement" in his program. Other comments by Thorndike include: (1) It should not be used for "the accurate diagnosis of reading, spelling, and arithmetic disabilities," as the authors suggest. (2) Procedures for determining validity and reliability data and norms are questionable. (3) Strictness in timing and in scoring could vary from examiner to examiner, and thus produce varying scores. Comments by users of the test include: 1. It can be used for students on any level, and it is not necessary to know which level the student should take before beginning to administer the test. 2. WRAT is useful as a screening test.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	No attempt was made to obtain a representative sampling. Such a sampling was not considered essential.
B.	How large was the sample?	5,933 (2,970 males and 2,963 females)
C.	From what geographic regions was the sample drawn?	Not specifically stated.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported?	Raw scores, grade equivalents and percentiles by age and subtest.
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Realiability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is the coefficient?	Ranges in .90+ for all three subtests.
C.	What type of reliability was reported?	Split-half.
D.	What was the composition of the group for whom reliability and coefficients were computed?	Groups were selected to represent probability distributions of achievements based on normative data.
E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	Ranges from 1.12 to 1.70 of raw scores for the three subtests depending upon age.

IV. Practicality:

A.	In administering the tests are time limits fixed or flexible?	Fixed for most parts.
B.	What is the total time needed to score the test?	Two to five minutes.
C.	Can the subtests be administered at different times?	Yes.

D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Thorough knowledge of the testing procedure is needed, particularly in pronunciation of words in spelling and reading subtests. A good basic understanding of tests and measurements is needed to interpret test results. A Buros reviewer, Jack C. Merwin, states that he finds the administration procedures confusing.
F.	Is it group administered?	Some parts are administered individually, others are group administered.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	Yes, manual is very complete in this respect.
I.	Are separate answer sheets available?	Student writes his answers in the four page test booklet.
J.	What is the range for the test series?	Pre-school through college.
K.	What is the origin of the test?	Was first standardized in 1936. It was designed as an adjunct to tests of intelligence and behavior adjustment.
L.	How recently has the test been revised?	1965.
M.	What were the special adaptations, if any, for adults?	Test was designed for adults and standardized on adult populations
N.	What is the background of the authors and publishers?	Authors and publishers have produced a mental ability test and books on mental retardation and measuring behavior.
O.	Is the format attrative and easy to follow?	Yes. Student actually writes words and performs arithmetic computations.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Four page test booklets and Manual of Instruction.

R.	What parts of the test are reusable and which must be repurchased?	Test booklets are consumed at each testing session.

Additional achievement tests that you may wish to obtain and review:

California Achievement Test (1963 edition of High School level and 1970 edition of all levels)

California Test Bureau
Del Monte Research Park
Monterey, California 93940

Comprehensive Tests of Basic Skills (4 levels)

California Test Bureau
Del Monte Research Park
Monterey, California 93940

Iowa Tests of Basic Skills

Houghton Mifflin
53 West 43rd Street
New York, New York 10036

Stanford Achievement Test

Harcourt Brace and Jovanovich, Inc.
757 Third Avenue
New York, New York 10017

Basic Reading Inventory

General Information:

1.	Title and level	Basic Reading Inventory
2.	Author	Richard W. Burnett
3.	Publisher	Scholastic Testing Service, Inc. 480 Meyer Road, Bensenville, Illinois 60106
4,	Date of Publication	1966
5.	Cost	$7.00/pkg. of 20 booklets
6.	Time for Administration	Less than one hour
7.	Number of forms of the test	One. Publisher reports no plans to expand test
8.	Type of test	Diagnostic
9.	Skill Range	0-5
10.	Sections, Parts, Subtests	Pt. 1) Sightwords Pt. 2) Sound and Letter Discrimination Beginning Consonants; Blends & Digraphs Pt. 3) Word Meaning (reading) Pt. 4) Word Meaning (listening) Pt. 5) Context Reading

I. Validity:

A.	What does the test measure (content)?	Pt. 1: Ability to underline the printed word out of 4 choices that corresponds to the given picture. Pt. 2: Ability to underline the one word out of a choice of four that begins with the same sound as the key word provided by teacher. Beginning consonants, blends and digraphs are included. Pt. 3: Ability to ascertain meaning from the printed word by underlining the correct synonym to an initial key word. Testee must recognize words independently. Words are graduated in difficulty.

	Pt. 4: Utilizes the same words as in part three, but tests the ability to underline the correct synonym when words are read to the testee.
	Pt. 5: Ability to read and comprehend short paragraphs of graded difficulty. Questions are of a factual and inferential type.
B. What type of validity does it exhibit?	Criterion.
1. From what sources were the test items drawn?	Complete information is not available. Paragraphs in Part 5 were adapted from The World Book Encyclopedia.
2. With what other tests or instruments was this test correlated?	The Gates Advanced Primary Reading Test. Teacher ratings.
What is the correlation coefficient??	Study 1: N = 128 pupils in grades 2 & 3 Gates total - BRI total r = .88 Teacher ratings - BRI total r =.76 Study 2: N = 40 fourth grade pupils. Gates total - BRI total r = .83 Teacher ratings - BRI total r =.84 Both studies were conducted in a northern Illinois industrial community.
What does the criterion test measure?	The Gates Advanced Primary Reading Test measures word recognition through picture clues and reading comprehension through questions referring to paragraphs.
3. What is the rationale for the test?	No information provided.
C. What skills does your program want to measure?	To be answered by program.
D. Examine the test itself.	
1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2. Are the test items well constructed?	To be answered by program.

3. Are they free from ambiguity?	To be answered by program.
E. What have reviewers, critics, and users of the test said about the test?	A review of this test by Albert J. Kingston in Buros' 7th MMY questioned the dearth of information regarding the test's rationale, reliability, and validity. The reviewer also noted that, since the test was published in 1965, more information on its usefulness with adults and norms based on adult groups should have been developed.
II. Standardization Sample:	No standardization sample.
A. What was the composition of the persons on whom the test was standardized?	
B. How large was the sample?	
C. From what geographic regions was the sample drawn?	
D. Are local norms available?	No.
E. In what terms are the norms reported? (grade level, age, etc.)	Raw score. Grade level scores can be inferred from results on Parts 3 and 5.
F. Are students similar to those in your program represented in the sample?	To be answered by program.
III. Reliability:	
A. Is a reliability coefficient reported?	Yes.
B. What is this coefficient?	.98 for total score. .97 for Part Three of the test.
C. What type of reliability coefficient was reported?	Test - re-test.
D. What was the composition of the group for whom reliability coefficients were computed?	.38 adult students enrolled in an urban adult literacy program in Northern Illinois.
E. Is a standard error of measurement reported?	No.

F. What is it?

IV. Practicality:

A.	In administering the tests are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the tests?	About 10-15 minutes.
C.	Can subdivisions of the test be administered at different times?	Parts 1-5 can be administered separately.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Knowledge of administration, scoring and interpretation of scores. Background in reading is very helpful.
F.	Is it group adminstered?	Yes. Groups of 20-30 students can be tested at one time.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	Yes. Chapter four provides information for discovering the students instructional level and for making instructional decisions based on test results.
I.	Are separate answer sheets available?	No.
J.	What is the range for the test series?	No series.
K.	What is the origin of the test?	Originally developed under the sponsorship of Adult Education Dept., State of Illinois.
L.	How recently has the test been revised?	First published in 1966.
M.	What were the special adaptations, if any, for adults?	Designed to assist in identifying functionally illiterate adolescents and adults.

N.	What is the background of the authors and publishers?	Unknown.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchases?	Manual and booklets, markers.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets must be repurchased.

Individual Reading Placement Inventory

General Information:

1.	Title and level	Individual Reading Placement Inventory
2.	Author	Edwin H. Smith and Weldon G. Bradtmueller
3.	Publisher	Follett Educational Corporation 1010 W. Washington Blvd., Chicago, Ill.
4.	Date of Publication	1969
5.	Cost	Pkg. of 20 Booklets, either form: $3.60. Administrator/teacher Packet: $4.50
6.	Time for Administration	Varies with each student
7.	Number of forms of the test	Two, A and B
8.	Type of test	Diagnostic
9.	Skill Range	0 - 7
10.	Sections, Parts, Subtests	Five: 1) Word Recognition and Analysis 2) Oral Reading Paragraphs 3) Present Language Potential 4) Auditory Discrimination 5) Letters of the Alphabet Part 1: Separate lists for levels 1.5 to 6.5 Part 2: Separate paragraphs for levels 1 to 6 Part 4: Initial sounds. Short vowel sounds, long vowel sounds, word endings, initial blends and digraphs

I. Validity:

A.	What does the test measure (content)?	Part 1: Ability to call words assumed to be on the grade levels of the specified list. The words have been checked with the Thorndike-Lorge and Taylor-Frackenpohl lists. Also tests the student's ability to analyze words not recognized on sight. Part 2: Ability to read orally and comprehend paragraphs of graded difficulty; or the ability to recognize

words in context and to exhibit comprehension of the relationships of those words by answering questions. The questions are primarily those of the factual recall and inference type.

Part 3: Ability to answer factual and inference questions about graded paragraphs read aloud to the student. Only those paragraphs not read orally by the student are read to him.

Part 4: Auditory discrimination of initial sounds, vowel sounds, ending sounds, and initial digraphs and blends. Students are asked to designate which of a list of several words is different from the others. There are four lists in each category.

Part 5: The ability to recognize and name the letters of the alphabet, and the ability to pronounce the sound(s) of those letters in isolation.

Also, Parts 1 and 2 provide information for determining the student's frustration, instructional, and independent reading levels.

B.	What type of validity does it exhibit?	Content, criterion (concurrent).
1.	From what sources were the test items drawn:	Words from the word recognition list and the graded paragraphs have been checked against their classifications on the Thorndike-Lorge and Taylor-Frackenpohl lists. The reading level of paragraphs has been validated by using the Spache and Dolch Readability for levels 1-3 and the Dale-Chall formula for levels 4-6. Parts 4 and 5 are to be administered to students who score below level 1. These skills have been assumed to be necessary ones to attain beginning reading ability. A large body of research supports this assumption.
2.	With what other tests or instruments was this test correlated?	ABE Student Survey by Rasof and Neff. The Stanford Reading Achievement Test. The California Reading Achievement Test. The Educational Development Series, Elementary and Advanced levels. Iowa Tests of Basic Skills.
	What is the correlation coefficent?	

	N	r
ABE Student Survey	146	.89
Stanford Achievement	75	.78
California Achievement	104	.87

Sections of IRPI were correlated to sections of EDS and the Iowa Tests of Basic Skills.

	N	EDS Basic Skills	EDS Reading
IRPI Word Recognition	108	.58	.69
IRPI Paragraph Reading		.48	.59

	N	Vocab.	Iowa Comp.	Ave. Rdg.
IRPI Wd. Recognition	143	.62	.62	.62
IRPI Paragraph Rdg.		.52	.47	.59

What does the criterion test measure?

See review of ABE Student Survey in this booklet.
See review of Tests of Adult Basic Education in this booklet.
The Stanford Achievement Test measures knowledge of word meaning in context, and paragraph comprehension through the correct insertion of words missing from the context of paragraphs. For the above correlation studies the authors do not specify how the reading score was correlated with scores on the criterion measures or the sections of the IRPI which were used for correlation. Correlations for different grade levels within the test range are not reported. The correlation studies for the above tests utilized adults in a migrant adult education program in Florida and junior and senior high school students in Northern Illinois.
The IRPI-EDS correlation study utilized 108 middle school students who had previously exhibited reading problems. In this study the testing was administered by undergraduate students in a reading methods course. The IRPI-Iowa correlation study utilized 143 sixth grade students. The Iowa test measures paragraph comprehension, noting details, organization and total meaning through paragraphs of varied length followed by multiple choice questions. The Vocabulary Section measures the ability to define words in context through a multiple choice format.

3. What is the rationale for the test?

None given in manual.

C. What skills does your program want to measure?

To be answered by program.

D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Reviewers in Buro's 7th MMY questioned the validity of the independent instructional and frustration levels, and call for further evidence of reliability and intercorrelation among parts of the test. Although criticizing the adequacy of the validity and standardization procedures, the reviewers stated that this test would be an aid in planning instruction for individuals. Reviewers of this test were Edward B. Pry and Albert J. Kingston.
II.	Standardization Sample:	Since grade level scores are determined through using the criteria of the frustration, instructional and independent levels, no norms are reported. The justification for the score lies in the accuracy of the reading level of the paragraphs, the rationale for frustration, instructional and independent levels, and the accuracy of the scorer.
A.	What was the composition of the persons on whom the test was standardized?	
B.	How large was the sample?	
C.	From what geographic regions was the sample drawn?	
D.	Are local norms available?	
E.	In what terms are the norms reported? (grade level, age, etc.)	
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Realiability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is this coefficient?	Study #1 - Ranges from .91 to .98; Study #2 - .66.
C.	What type of reliability coefficient was reported?	(1) & (2) - Equivalent form. (2) also indicates scorer reliability as different scorers administered the forms.
D.	What was the composition of the group for whom reliability coefficients were computed?	(1) 410 students from ABE classes in Florida, ABE classes in a Florida prison, junior and senior high school retarded readers in Florida, North Illinois, and South Carolina (2) 25 eighth grade students.
E.	Is a standard error of measurement reported?	No.
F.	What is it?	

IV. Practicality:

A.	In administering the tests are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	Varies with each administration depending upon sections given.
C.	Can the subtests be administered at different times?	Yes, but not desirable.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	A thorough knowledge of the test procedure and scoring. A background in the administration of oral reading tests and in basic knowledge of reading.
F.	Is it group administered?	Individually administered.
G.	Are the directions easily followed by students?	Yes.

H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	Separate answer booklets are necessary for each student.
J.	What is the range for the test series?	0-7.
K.	What is the origin of the test?	Authors report that the test was used widely in Florida prior to publication.
L.	How recently has the test been revised?	Initial publication 1969.
M.	What were the special adaptations, if any for adults?	Test specially designed for adolescents and adults. Scoring of errors for oral reading paragraphs considers speech patterns of disadvantaged students.
N.	What is the background of the authors and publishers?	Edwin Smith is a well-known reading specialist particularly interested in adolescents and adults. He has written in conjunction with others the Reading Development Kit, Literacy Ed. for Adoles. and Adults and the Sound Spelling Program. Dr. Weldon Bradtmueller is Associate Professor of Education at Northern Illinois University. Follett Ed. Corp. publishes many instructional materials for remedial adolescents and adults, and one other test, the ABE Student Survey.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	A manual, Form A & B word wheels, Form A & B reading paragraph cards, and multiple copies of the test and scoring booklets for Forms A & B.
R.	What parts of the test are reusable and which must be repurchased?	A copy of the Student Test and Scoring Booklet is consumed per each administration of the test.

Additional tests which may be helpful in diagnosing adult students' abilities:

Botel Reading Inventory Test

Follett Publishing Co.
1010 W. Washington Blvd.
Chicago, Ill. 60607

Prescriptive Mathematics Inventory

California Test Bureau
Del Monte Research Park
Monterey, California 93940

English-Second-Language Placement Test, 100-200-300

General Information:

1.	Title and level	English-Second-Language Placement Test, 100-200-300
2.	Author	Donna Ilyin
3.	Publisher	San Francisco Community College District Alemany Adult School 750 Eddy Street San Francisco, California 94109
4.	Date of Publication	1969
5.	Cost	10¢ each. Kit including examination copy of each form is available free of charge
6.	Time for Administration	30 Minutes
7.	Number of forms of the test.	Two: Forms A and B
8.	Type of test	Screening and achievement
9.	Skill Range	First 3 (out of a possible 6) levels of ESL
10.	Sections, Parts and Subtests.	None

I. Validity:

A.	What does the test measure (content)?	Ability to read English sentences and choose the ones that use correct structures.
B.	What type of validity does it exhibit?	Content, criterion, construct.
	1. From what sources were the test items drawn?	Based on curriculum content of San Francisco Community College District Adult Schools.
	2. With what other tests or instruments was this test correlated?	Test of Aural Comprehension, Form C (picture section only) by Robert Lado: Aural Comprehension Tests and Grammar Contrast Tests by Paul Nixon; and Ilyin Oral Interview by Donna Ilyin.
	What is the correlation coefficient?	.72, .80, and .85.

	What does the criterion test measure?	Lado and Nixon: Listening comprehension; answers given by indicating a picture that relates to a question or statement. Ilyin: Listening and speaking skills.
	3. What is the rationale for the test?	Rationale for a structure test is that most ESL learning materials are built around structural patterns presented sequentially. The test was designed for placement of students in an adult school ESL curriculum, rather than in a college curriculum.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers and critics and users of the test said about the test?	Not available at time of review.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	Adult evening school and full-time day adult students in the cities of Sacramento and San Francisco, migrant workers enrolled in ESL classes, and college students of ESL.
B.	How large was the sample?	1600.
C.	From what geographic regions was the sample drawn?	Urban and rural areas of California.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported? (grade level, age, etc.)	Class level, expressed as 100,200,300, 400, 500, 600. (See ESL Master Plan, Phase II, available from Dr. Steven Morena, Assistant Superintendent, San Francisco Community College District, 33 Gough Street, San Francisco, Calif.)

F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is this coefficient?	Form correlation: .93, Form A: .91, Form B: .92, Overall: .95.
C.	What type of reliability coefficient was reported?	Alternate form; Internal consistency.
D.	What was the composition of the group for whom reliability coefficients were computed?	100-200-300 level students (heterogeneous language backgrounds) attending adult schools in San Francisco.
E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	Form A: 3.02; Form B: 3.06.

IV. Practicality:

A.	In administering the tests are time limits fixed or flexible?	Fixed.
B.	What is t[illegible] needed to s[illegible] test?	2-3 minutes.
C.	Can the subtests be administered at different times?	N.A.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	None.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Yes.

H.	Does the manual contain guides for using and interpreting the scores?	Instruction sheet with norms and books used at each level is available. Manual will be available in September, 1972.
I.	Are separate anwer sheets available?	Yes, but not needed.
J.	What is the range for the test series?	Two test levels, 100-200-300 and 400-500-600, place students in six levels of adult school ESL classes.
K.	What is the origin of the test?	Sampling of structures taught in 100, 200 and 300 level classes in San Francisco adult schools.
L.	How recently has the test been revised?	1972. Revised forms (C and D) will be available soon.
M.	What were the special adaptations, if any, for adults?	Designed for adults.
N.	What is the background of the authors and publishers?	Donna Ilyin is Adult Classroom Teacher and Consultant for Testing and Teacher Training in the San Francisco Community College District and has served as a consultant for the California State Department of Education, the University of San Francisco, the University of Southern California, San Francisco State College, Arizona State University, and Sonoma State College.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	One copy of booklets for Forms A and B can be obtained free of charge. Additional copies may be purchased; however, programs may reproduce copies for their own use, provided that proper credit be given.
R.	What parts of the test are reusable and which must be repurchased?	Additional test booklets must be purchased or reproduced.

English-Second-Language Placement Test, 400-500-600

General Information:

1.	Title and level	English-Second-Language Placement Test, 400-500-600
2.	Author	Donna Ilyin, Jeanette Best, and Virginia Biagi
3.	Publisher	San Francisco Community College District Alemany Adult School 750 Eddy Street San Francisco, California 94109
4.	Date of Publication	1972
5.	Cost	Not available at time of review
6.	Time for Administration	30 minutes
7.	Number of forms of the test	Two: Forms G and H
8.	Type of test	Screening and achievement
9.	Skill Range	Last 3 (out of a possible 6) levels of ESL
10.	Sections, Parts, Subtests	None

I. Validity:

A.	What does the test measure (content)?	Ability to read English sentences and choose the ones that use correct structures.
B.	What type of validity does it exhibit?	Content, criterion, and construct.
	1. From what sources were the test items drawn?	Based on curriculum content of San Francisco Community College District adult schools.
	2. With what other tests or instruments was this test correlated?	Long Beach City College Test, by Donald Mills; Comprehensive English Language Test for Speakers of English as a Second Language (CHLT), Structure Test by David P. Harris and Leslie A. Palmer; and An English Reading Test, by Harold V. King and Russell N. Campbell.
	What is the correlation coefficient?	Not available at time of review.

	What does the criterion test measure?	Mills: Minimal-pair listening and choosing correct structures for sentences read by testee. Harris and Palmer: Ability to read and select acceptable constructions in sentences. King and Campbell: Knowledge of word and idiom meanings in context and ability to answer multiple choice items based on paragraphs read by testee.
3.	What is the rationale for the test?	Rationale for a structure test is that most ESL learning materials are built around structural patterns presented sequentially. The test was designed for placement of students in an adult school curriculum, rather than in a college curriculum.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers and critics and users of the test said about the test?	Not available at time of review.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	Adults enrolled in evening and full-time day adult schools, in special vocational programs, and in college preparatory programs, all in San Francisco.
B.	How large was the sample?	1200.
C.	From what geographic regions was the sample drawn?	San Francisco.
D.	Are local norms available?	To be answered by program.

E.	In what terms are the norms reported? (grade level, age, etc.)	Class level, expressed as 100, 200, 300, 400, 500, 600. (See ESL Master Plan, Phase II, available from Dr. Steven Morena, Asst. Superintendent, San Francisco Community College District, 33 Gough Street, San Francisco.)
F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is this coefficient?	Form correlation: .86; Form G: .89, Form H: .88.
C.	What type of reliability coefficient was reported?	Alternate form; Internal consistency.
D.	What was the composition of the group for whom reliability coefficients were computed?	400-500-600 level students (Heterogeneous language backgrounds) attending San Francisco adult schools.
E.	Is a standard error of measurement reported?	Not available at time of review.
F.	What is it?	Not available at time of review.

IV. Practicality:

A.	In administering the tests are time limits fixed or flexible?	Fixed.
B.	What is the total time needed to score the test?	2-3 minutes.
C.	Can the subtests be administered at different times?	N.A.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	None.

F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Yes.
H.	Does the manual contain guides for using and interpreting the scores?	Manual will be available in September, 1972.
I.	Are separate answer sheets available?	Yes, but not needed.
J.	What is the range for the test series?	Two test levels, 100-200-300 and 400-500-600, place students in six levels of adult school ESL classes.
K.	What is the origin of the test?	Sampling of structures taught in 400,500 and 600 level classes in San Francisco adult schools.
L.	How recently has the test been revised?	Test to be published in September, 1972.
M.	What were the special adaptations, if any, for adults?	Designed for adults.
N.	What is the background of the authors and publishers?	The three authors are Adult Classroom Teachers and consultants for the San Francisco Community College District. Donna Ilyin has served as a consultant for the California State Department of Education and for several colleges and universities in the California area. Jeannette Best has been a consultant to the California State Department of Education, and chairman of a State Department Vocational Curriculum Project. Virginia Biagi served as a bilingual coordinator for the San Francisco Unified School District.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets.

R.	What parts of the test are reusable and which must be repurchased?	Not available at time of review.

Examination in Structure

General Information:

1.	Title and level	Examination in Structure
2.	Author	Charles C. Fries and Robert Lado
3.	Publisher	English Language Institute, University of Michigan. May be purchased from: Follett's Michigan Book Store 322 South State Street Ann Arbor, Michigan 48108
4.	Date of Publication	Copyright 1947; Reprinted 1955
5.	Cost	Testing Package (20 test booklets, 100; answer sheets, scoring stencil) - $6.00; Booklets - pkg. of 20 - $3.00; Answer Sheets - pkg of 100 - $3.00; Scoring Stencil - $1.50; Specimen Set - $1.50
6.	Time for Administration	60 Minutes
7.	Number of forms of the test	Three: Forms A, B, and C
8.		
8.	Type of test	Achievement
9.	Skill Range	Intermediate - Advanced
10.	Section, Parts, Sub-tests	Eleven parts

I. Validity:

A.	What does the test measure (content)?	Ability to read a sentence and choose one of two or three words or phrases to make it grammatically correct. Includes verbs, prepositions, adverbs, pronouns, and other parts of a speech. Ability to provide correct pronouns, plurals, interrogatives, negatives, forms of verbs, and sentence order by filling in blanks.
B.	What type of validity does it exhibit?	Information on validity not available.
	1. From what sources were the test items drawn?	

2.	With what other tests or instruments was this test correlated?	
	What is the correlation coefficient?	
3.	What is the rationale for the test?	
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers and critics and users of the test said about the test?	Information not available at time of review.
II.	Standardization Sample:	(Norms not available).
A.	What was the composition of the persons on whom the test was standardized?	
B.	How large was the sample?	
C.	From what geographic regions was the sample drawn?	
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported? (grade level, age, etc.)	
F.	Are students similar to those in your program represented in the sample?	

III.	Reliability:	Information on reliability not available.
A.	Is a reliability coefficient reported?	No.
B.	What is this coefficient?	
C.	What type of reliability coefficient was reported?	
D.	What was the composition of the group for whom reliability coefficients were computed?	
E.	Is a standard error of measurement reported?	
F.	What is it?	
IV.	Practicality:	
A.	In administering the tests are time limits fixed or flexible?	Fixed (Since there are no norms, the 60 minutes time limit need not be adhered to).
B.	What is the total time needed to score the test?	About 5 minutes.
C.	Can the subtests be administered at different times?	Yes.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	None.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	If students can read English on the level of the test items, should not have difficulty following the directions.
H.	Does the manual contain guides for using and interpreting the scores?	No manual available.

I.	Are separate answer sheets available?	Yes.
J.	What is the range for the test series?	Intermediate - Advanced.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	No revision.
M.	What were the special adaptations, if any, for adults?	Written for adults.
N.	What is the background of the authors and publishers?	Both Robert Lado and Charles C. Fries have been directors of the English Language Institute, University of Michigan. Lado is presently Dean of the School of Linguisties, Georgetown University. Both Fries, now deceased, and Lado have authored numerous materials on ESL, including ESL textbooks and tests, and books and articles for training teachers and professional reading.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Test booklets, answer sheets, scoring stencil.
R.	What parts of the test are reusable and which must be repurchased?	Test booklets are reusable; answer sheets must be repurchased.

Ilyin Oral Interview

General Information:

1.	Title and level	Ilyin Oral Interview, Experimental Edition
2.	Author	Donna Ilyin; Artist: Tanya Butler
3.	Publisher	Newbury House Publishers, 68 Middle Rd., Rowley, Massachusetts 01969
4.	Date of Publication	1972
5.	Cost	Test Book $10.00; Manual $1.25; Pad of Answer Sheets $1.75
6.	Time for Administration	5-30 minutes for each student tested. The test is terminated at the frustration level
7.	Number of forms of the test	Two: BILL and SAM
8.	Type of test	Screening or Achievement
9.	Skill Range	First 5 (out of a possible 6) levels of ESL
10.	Sections, Parts, Subtests	None

I. Validity:

A.	What does the test measure (content)?	Ability to understand and speak original English sentences giving correct information and using acceptable word order, verb form, and other structures.
B.	What type of validity does it exhibit?	This is an experimental edition of the test and validity data is still being collected. The author and publishers will welcome information from programs which can provide comparative scores on other standardized tests.
	1. From what sources were the test items drawn?	Unknown.
	2. With what other tests or instruments was this test correlated?	Not available at time of review.

	What is the correlation coefficient?	Not available at time of review.
	What does the criterion test measure?	Not available at time of review.
3.	What is the rationale for the test?	Many students learning English or another new language are not tested in any standardized way for their ability to communicate through understanding and speaking the language even though modern methods of teaching emphasize these skills. And some people learn to communicate and conduct daily affairs in the new language and yet are unable to read and write it or to speak an educated form of it. The Ilyin Oral Interview is designed to test such individuals' ability to use English orally in response to hearing it, in a controlled situation, requiring the student to do no reading or writing. Often oral interviews are quite time-consuming, as each interview is graded by two or three examiners who use a subjective ranking scale. These subjective scores are then averaged to get a fairer grade for each individual tested. Since the Ilyin Oral Interview has a more objective system of grading allowing greater consistency, only one examiner is needed for each candidate tested.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers and critics and users of the test said about the test?	Not available at time of review.

II. Standardization Sample:

A.	What was the composition of the persons on whom the test was standardized?	The student populations tested in developing the interview included 150 students from many native language backgrounds in day adult ESL classes in a large urban area. The students ranged in age from 18 to 63, and came from Asian, Latin American, European and Middle Eastern countries. Some of them worked with native speakers, while others had little contact with oral English except in class. Another test group consisted of 300 Spanish-speaking students enrolled in vocational programs in large urban areas. Their ages ranged from 18 to 45; their educational backgrounds ranged from the fourth grade level through the completion of university work. These students had lived in the country from six months to twenty years, and had had varying amounts of contact with native speakers of English. The experimental edition is being tested further on adult students in regular adult day programs and in vocational training projects. It is also being tested on junior high school students in a Spanish bi-lingual program, and will be used with prospective college students. Results of the testing, norms, and reliability will be forthcoming. Reports from other programs using the experimental edition of the interview will be welcomed by the author and publishers.
B.	How large was the sample?	Not available at time of review.
C.	From what geographic regions was the sample drawn?	Urban California.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported? (grade level, age, etc.)	Norms are still being developed. Preliminary norms are suggested in the manual, placing students into 5 (out of a possible 6) levels of ESL. (See ESL Master Plan Phase II, available from Dr. Steven Morena, Assistant Superintendent, San Francisco Community College District, 33 Gough Street, San Francisco, California.)

F.	Are students similar to those in your program represented in the sample?	To be answered by program.

III. Reliability:

A.	Is a reliability coefficient reported?	Data is still being collected.
B.	What is this coefficient?	Not available at time of review.
C.	What type of reliability coefficient was reported?	Not available at time of review.
D.	What was the composition of the group for whom reliability coefficients were computed?	See answer to II.A. (above).
E.	Is a standard error of measurement reported?	Not available at time of review.
F.	What is it?	Not available at time of review.

IV. Practicality:

A.	In administering the tests are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	5-10 minutes. The manual suggests a shorter, more subjective manner of scoring.
C.	Can the subtests be administered at different times?	Not available at time of review.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	The examiner needs to: -speak in a normal conversational manner using speech patterns that will be familiar to the individuals tested; -recognize if a response reflects correct information, correct word order and correct verb structures, and know if other structures are correct; -be able to transcribe the exact response that is made.

F.	Is it group administered?	No, individually.
G.	Are the directions easily followed by students?	Yes .
H.	Does the manual contain guides for using and interpreting the scores?	Complete information cannot be made available until research testing is completed. The manual, however, is very complete in explaining how to administer and score the test, and interpret individual responses.
I.	Are separate answer sheets available?	Yes.
J.	What is the range for the test series?	Beginning through Advanced.
K.	What is the origin of the test?	Based on teachers' experience in teaching ESL to adults.
L.	How recently has the test been revised?	Experimental edition published in 1972.
M.	What were the special adaptations, if any, for adults?	Designed especially for adults.
N.	What is the background of the authors and publishers?	Donna Ilyin is Adult Classroom teacher and Consultant for Testing and Teacher Training in the San Francisco Community College District and has served as a consultant for the California State Department of Education, the University of San Francisco, the University of Southern California, San Francisco State College, Arizona State University, and Sonoma State College. The artist, Tanya Butler, is an immigrant from the Soviet Union, and a well-known portrait artist and cartoonist in the San Francisco area.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	No reading required.
Q.	What are the components to be initially purchased?	Test Book, Manual, and Answer Sheets.
R.	What parts of the tests are reusable and which must be repurchased?	Answer sheets must be repurchased.

Oral Placement Test for Adults (SWCEL)

General Information:

1.	Title and level	Oral Placement Test for Adults, Experimental Edition (SWCEL)
2.	Author	Allen Ferrel
3.	Publisher	Southwest Cooperative Lab. Inc. 117 Richard Drive N.E. Albuquerque, New Mexico 87106
4.	Date of Publication	June, 1971
5.	Cost	$.50/booklet
6.	Time for Administration	No time limits. Depends on student level. Maximum: 2-8 minutes/person
7.	Number of forms of the test	One
8.	Type of test	Screening
9.	Skill Range	Complete inability to speak and understand oral English to a level where the student can understand oral English and can participate in normal conversation with few errors
10.	Sections, Parts, Subtests	4 Subtests

I. Validity:

A. What does the test measure (content)?

Part 1: Students ability to verbally respond effectively to 7 questions asked by the interviewer. Correct English is not required.

Part 2: Ability to repeat (with understandable pronunciation) 12 out of 14 items which are initally pronounced by the interviewer.

Part 3: Items 1-5: Ability to orally respond to personal questions posed by the interviewer.

Items 6-10: Ability to transpose a positive statement into a negative statement. Pronoun and verb must agree in the student reply.

		Items 11-15: Ability to transpose a statement to a question. Student must respond correctly to all 15 items.
		Part 4: Ability to change present tense statements to simple past tense; ability to change items into future tense. Students must answer all ten items correctly to complete level 4 successfully.
B.	What type of validity does it exhibit?	Construct, criterion.
	1. From what sources were the test items drawn?	Everyday situational statements; Typical entry form questions.
	2. With what other tests or instruments was this test correlated?	Teacher placement based on subject interview.
	What is the correlation coefficient?	Not available at time of review.
	3. What is the rationale for the test?	That measurement of ESL students should be based on their functional use of English.
C.	What skills does your program measure?	To be answered by program.
	What does the criterion test measure	Subjective placement of students by teachers
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Not available at time of review.

II.	Standarization Sample:	None reported.
A.	What was the composition of the persons on whom the test was standardized?	
B.	How large was the sample?	
C.	From what geographic regions was the sample drawn?	
D.	Are local norms available?	
E.	In what terms are the norms reported?	Proficiency scale for placement of students is included.
F.	Are students similar to those in your program represented in the sample?	
III.	Reliability:	Not reported.
A.	Is a reliability coefficient reported?	
B.	What is the coefficient?	
C.	What type of reliability was reported?	
D.	What was the composition of the group for whom reliability coefficients were computed?	
E.	Is a standard error of measurement reported?	
F.	What is it?	
IV.	Practicality:	
A.	In administering the tests are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	Can be scored as it is administered.

C.	Can the subtests be administered at different times?	Yes. Not advisable.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Practice in listening to and scoring ESL students.
F.	Is it group administered?	Individually administered.
G.	Are the directions easily followed by students?	Directions become more complex as tasks become more difficult.
H.	Does the manual contain guides for using and interpreting the scores?	Yes.
I.	Are separate answer sheets available?	No, only booklets.
J.	What is the range for the test series?	Not a series.
K.	What is the origin of the test?	Developed under U.S. Office of Education Special Project Grant.
L.	How recently has the test been revised?	Initial edition. No plans for revision.
M.	What were the special adaptations, if any, for adults?	Especially constructed for adults.
N.	What is the background of the authors and publishers?	Publisher is a specially funded government project.
O.	Is the format attractive and easy to follow?	Student does not see test.
P.	Is the print size appropriate?	Student does not see test.
Q.	What are the components to be initially purchased?	Test booklet.
R.	What parts of the test are reusable and which must be repurchased?	If administered as directed, each testing session would consume a booklet. However, separate answer sheets could be typed and duplicated by a program.

Test of Aural Comprehension

General Information:

1.	Title and level	Test of Aural Comprehension
2.	Author	Robert Lado
3.	Publisher	English Language Institute, University of Michigan
4.	Date of publication	1957
5.	Cost	Test Booklets: $3.00/20; Answer Sheets: $2.00/100; Examiner's Book: $1.50; Entire package: $6.00
6.	Time for administration	40 minutes
7.	Number of forms of the test	Forms A, B, C
8.	Type of test	Screening or Achievement
9.	Skill range	Beginners to Advanced (University Level)
10.	Sections, parts, subtests	Two parts

I. Validity:

A. What does the test measure (content)?	Part 1:	Examiner reads sentence to student(s), who must choose one of three pictures that best illustrates the aural statement. Task requires student ability to identify picture that matches a verb tense, a direct question, the entire sentence, a plural cue, and a possessive cue; also, to auditorially distinguish between similar sounds.
	Part 2:	Student selects one of three (written) answers that answers an orally-provided statement. Requires ability to understand oral cue, to know an appropriate response, to read selections, and then to select the correct answer.

B.	What type of validity does it exhibit?	Construct, criterion.
	1. From what sources were the test items drawn?	Unknown.
	2. With what other tests or instruments was this test correlated?	Teacher Opinion (8 teachers).
	What is the correlation coefficient?	.85.
	3. What is the rationale for the test?	"That it tests student understanding of spoken sentences and paragraphs, and those sentences and paragraphs have been especially constructed to bring out the comprehension problems of the language."
C.	What skills does your program measure?	To be answered by program.
	What does the criterion test measure?	Subjective impression of student ability through daily contact.
D.	Examine the test itself.	
	1. Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
	2. Are the test items well constructed?	To be answered by program.
	3. Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Herschel T. Manual and Clarence E. Turner, who reviewed this test in Buro's 6th Mental Measurements Yearbook, found many positive features. They commented that the test would be useful in judging students' ability to work in English speaking colleges and that the student's ability to understand phrases that have proven to be difficult for non-English natives would be revealed. The reviewers did question the validity of specific items, several factors

		of test construction, the use of percentage scores, and the applicability of the norms to other settings.
II.	Standardization Sample:	Two sets of norms: Proficiency; Progress norms reflect progress of students who received 25 hours per week of ESL instruction.
A.	What was the composition of the persons on whom the test was standardized?	Unspecified. Some were students at the English Language Institute.
B.	How large was the sample?	500 for Proficiency; 1,000 for Progress.
C.	From what geographic regions was the sample drawn?	Unspecified.
D.	Are local norms available?	No.
E.	In what terms are the norms reported?	Percents (not percentiles). Raw scores.
F.	Are students similar to those in your program represented in the sample?	Test description indicates that norms are most appropriate for students contemplating academic study at a univeristy.
III.	Reliability:	
A.	Is a reliability coefficient reported?	Yes.
B.	What is the coefficient?	.87 between Forms A and C; .88, split-half.
C.	What type of reliability was reported?	Alternate form, split-half.
D.	What was the composition of the group for whom reliability and coefficients were computed?	Students at English Language Institutes. Further information unspecified. Reviewer inferred that students' goal was academic study at an English-speaking university.
E.	Is a standard error of measurement reported?	Yes.

F.	What is it?	.03 for alternate form; .02 for split-half.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Somewhat flexible.
B.	What is the total time needed to score the test?	About 2 minutes.
C.	Can the subtests be administered at different times?	Yes. Not advisable.
D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	No special requirements. Review of test description and instructions, and a reasonably clear voice.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	Directions on test booklet are printed in English, Spanish, Portuguese and French. Examples are provided to clarify task.
H.	Does the manual contain guides for using and interpreting the scores?	Proficiency norms and Progress norms.
I.	Are separate answer sheets available?	Yes.
J.	What is the range for the test series?	Only one test for entire range.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	Has not been revised.
M.	What were the special adaptations, if any, for adults?	Test seems to have been designed for educated ESL adults.

N.	What is the background of the authors and publishers?	Author and publisher are prominent in the field of ESL theory and materials.
O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Examiner's booklet, several copies of the test booklet, answer sheets, scoring stencil.
R.	What parts of the test are reusable and which must be repurchased?	Answer sheets must be repurchased.

Test of Aural Perception in English for Latin-American Students

General Information:

1.	Title and level	Test of Aural Perception in English for Latin-American Students
2.	Author	Robert Lado
3.	Publisher	English Language Institute, Univ. of Michigan. Can be purchased from: Follett's Michigan Book Store 322 South State St. Ann Arbor, Michigan 48108
4.	Date of publication	1957
5.	Cost	Examiner's Booklet: $1.50; Answer Sheets: $5.00/100; Scoring Stencil: $2.50
6.	Time for administration	About 50 minutes
7.	Number of forms of the test	One
8.	Type of test	Diagnosis of Auditory Perception Skills
9.	Skill range	Intermediate--Advanced
10.	Sections, parts, subtests	Part I: A and B; Part II: A and B

I. Validity:

A.	What does the test measure (content)?	Ability to hear which of three or four orally read words or sentences, which may differ by a single sound, are the same.
B.	What type of validity does it exhibit?	Construct.
	1. From what sources were the test items drawn?	Auditory perception problems of Spanish-speakers were chosen for the test.
	2. With what other tests or instruments was this test correlated?	None.

	What is the correlation coefficient?	None.
	What does the criterion test measure?	None.
3.	What is the rationale for the test?	Since pronunciation clues are isolated in the test, it is more valid as a measure of auditory perception than other means of observation. When given to 33 native speakers of English, their scores clustered at the top of the scale (93-100), with the number of cases increasing as the score approached 100. It was concluded, therefore, that native speakers of English hear the sound differences tested without difficulty. If the ESL student hears the significant sound differences in his native language, this test is valid in determining his auditory discrimination of the sound differences of English.
C.	What skills does your program want to measure?	To be answered by program.
D.	Examine the test itself.	
1.	Do the test items appear appropriate for measuring those abilities you want to measure?	To be answered by program.
2.	Are the test items well constructed?	To be answered by program.
3.	Are they free from ambiguity?	To be answered by program.
E.	What have reviewers, critics and users of the test said about the test?	Not available at time of review.
II.	Standardization Sample:	
A.	What was the composition of the persons on whom the test was standardized?	Students who completed eight weeks of intensive work in English at the English Language Institute, University of Michigan.

B.	How large was the sample?	406.
C.	From what geographic regions was the sample drawn?	Not reported.
D.	Are local norms available?	To be answered by program.
E.	In what terms are the norms reported (grade level, age, etc.)?	Raw scores (percents) are converted to deciles.
F.	Are students similar to those in your program represented in the sample?	No. The decile norms are of little value for ABE programs.

III. Reliability:

A.	Is a reliability coefficient reported?	Yes.
B.	What is this coefficient?	.90 for 75 items; .92 estimated for 100 items.
C.	What type of reliability coefficient was reported?	Test-retest.
D.	What was the composition of the group for whom reliability coefficients were computed?	Students at the English Language Institute, University of Michigan.
E.	Is a standard error of measurement reported?	Yes.
F.	What is it?	.02 for 75 items.

IV. Practicality:

A.	In administering the tests, are time limits fixed or flexible?	Flexible.
B.	What is the total time needed to score the test?	15 minutes.
C.	Can the subtests be administered at different times?	Yes.

D.	Will the test consume a reasonable portion of program/student time?	To be answered by program.
E.	What qualifications are needed to administer the test?	Should be administered by a native speaker of standard American English or by one who can speak English with all the pronunciation contrasts of a native speaker. Completion and interpretation of the diagnostic sheet requires knowledge of the International Phonetic Alphabet.
F.	Is it group administered?	Yes.
G.	Are the directions easily followed by students?	May be confusing for some, but sufficient examples should clarify difficulties.
H.	Does the manual contain guides for using and interpreting the scores?	Some.
I.	Are separate answer sheets available?	Yes.
J.	What is the range for the test series?	Intermediate--Advanced.
K.	What is the origin of the test?	Unknown.
L.	How recently has the test been revised?	No revision.
M.	What were the special adaptations, if any, for adults?	Designed for college students and adults.
N.	What is the background of the authors and publishers?	Robert Lado has been director of the English Language Institute, University of Michigan, which publishes numerous materials on ESL. Lado is presently Dean of the School of Linguistics, Georgetown University. He has authored many materials on ESL, including ESL textbooks and tests, and books and articles for training teachers and professional reading.

O.	Is the format attractive and easy to follow?	Yes.
P.	Is the print size appropriate?	Yes.
Q.	What are the components to be initially purchased?	Examiner's book, Answer sheets, scoring stencil.
R.	What parts of the test are reusable and which must be repurchased?	Answer sheets must be repurchased.

Additional tests which may prove useful in ESL Programs:

Tests In English

ESL Placement Test

by Donald Mills
Long Beach Community College
Long Beach City College Dist.
4901 E. Carson Street
Long Beach, Calif. 90808

ESL Proficiency Test

by Louis S. Marano
839-58th Street
Brooklyn, New York 11220

Michigan Test of English Language Proficiency

Follett's Michigan Bookstore
322 South State Street
Ann Arbor, Michigan 48108

Sample Placement Test for ESL

by Jean Bodman
1972 Summer ESL Institute
Available from the Adult
Education Resource Center
Jersey City State College
John F. Kennedy Blvd.
Jersey City, N.J.

Test in Speaking English

by Donald Ford
Available from Valerie M.
Comives
5057 Woodward
Detroit, Mich. 48202

Oral Production Tests (Levels 1, 2,3); Oral Placement Test (for use with Orientation in American English series)

Robert Poczik
Bureau of Basic Continuing
Education
State Dept. of Education
Albany, New York

Placement Examination (for use with Orientation in American English)

Institute of Modern
Language, Inc.
61 West 51st St.
New York, N.Y. 10009

Placement Tests for Speakers of Other Languages

by Alice Perlman
Staff Development Team
Bd. of Education of the
City of New York
130 Clinton St.
Brooklyn, N.Y. 11201

Vocational Language Skills Test

by Anne Terrell
Chinatown--No. Beach
English Language Center
550 Montgomery St.,
10th Floor
San Francisco, Calif. 94111

Tests In Spanish

CIA: Natural Sciences: Vocabulario E Interpretacion de Material de Lectura (grades 8-13) (1950 ed.)
CIA: Social Studies: Vocabulario E Interpretacion de Material de Lectura

Guidance Testing Associates
6516 Shirley Ave.
Austin, Texas 78752

Individual Test in Speaking Spanish

Staff Development Team
Adult Basic Education Program
130 Clinton Street
Brooklyn, New York 11201

Summary of Test Information

Test	No. of Forms	Type	Skill Range	Total Administration Time	Subject Areas Tested	Publication Date
ABE Student Survey	2	Achievement	0-8	Approx. 3 hrs.	Reading comprehension Word recognition Arithmetic computations and problems	1966
ABLE	2	Achievement	1-12	I & II--2 hrs. 15 mins. III--3 1/2 hrs.	Vocabulary Reading Spelling Arithmetic	I & II--1967 III--1971
Basic Reading Inventory	1	Diagnostic	0-5	less than 1 hr.	Sight words Auditory discrimination Word meaning Context Reading	1966
Dolch Basic Sight Word Test	1	Sight Vocabulary	0-3	5 mins.	-	1942
Fundamental Achievement	2	Achievement	literacy to 8 level	1 hr.	Verbal ability Numerical ability	1968
Harris Graded Word List	1	Sight Vocabulary	0-5	5 mins.	-	-
IRPI	2	Diagnostic	0-7	Varies	Word recognition and analysis Oral rdg. paragraphs Present lang. potential Auditory discrimination Letters of the alphabet	1969
ITED	2	Achievement	9-12	9 hrs. (1 hrs. for each test)	Understanding of Basic Social Concepts	1963

-mo

Summary of Test Information Continued

Test	No. of Forms	Type	Skill Range	Total Administration Time	Subject Areas Tested	Publication Date
					General Background in the Natural Sciences Correctness & Appropriateness of Expression Ability to Do Quantitative Thinking Ability to Interpret Reading materials in the Social Sciences Ability to Interpret Reading materials in the Natural Sciences Ability to Interpret Literary Materials General Vocabulary Uses of Sources of Information	
MAT	Elem.: 3/Int. Adv.: 4 H.S.: 1	Achievement	3-13	From 2 hrs. 27 mins. to 5 hrs. 15 mins.	Vocabulary, Reading Spelling, Language, Study Skills, Arithmetic Computation and Problem Solving, Social Studies, Science	From 1958 to 1963
Slosson Oral	1	Sight Vocabulary	0-HS	10 mins.	-	1963
SRA Reading & Arithmetic Indexes	1	Screening	1-8	25/index	Reading & Arithmetic	1968

-more-

Summary of Test Information Continued

Test	No. of Forms	Type	Skill Range	Total Administration Time	Subject Areas Tested	Publication Date
TABE	2	Achievement	2-9	E--1 1/2 hrs. M--2 1/2 hrs. D--3 hrs.	E--Rdg. & Arithmetic M & D--Rdg., Arith., Language	1967
Wide Range Achievement Test	1	Achievement	K thru college	20 to 30 mins.	Reading Spelling Arithmetic	1965 revised

Summary of Test Information

ESL Test	No. of Forms	Type	Skill Range	Total Administration Time	Subject Areas Tested	Publication Date
ESL Placement Test 100-300	2	ESL Screening & Achievement	First 3 out of 6 levels	30 mins./test	Reading English Sentences Selecting English Structures	1969
ESL Placement Test 400-600	2	ESL Screening & Achievement	Last 3 out of 6 levels	30 mins./test	Reading English Sentences Selecting English Structures	1972
Examination in Structure	3	ESL Achievement	Intermediate to Advanced	60 mins.	English Sentence Structure Reading	1947
Ilyin Oral Interview	2	ESL Screening or Achievement	First 5 out of 6 levels	5-30 mins.	Oral Communication in English	1972
SWCEL	1	Screening	Beginners to Intermediate	2-8 mins.	Ability to speak and Understand English	1971
Test of Aural Comprehension	3	Screening or Achievement	Beginners to Adavnced	40 mins.	Ability to speak and Understand English	1957
Test of Aural Perception in English for Latin-American Students	1	Diagnostic	Intermediate to Advanced	50 mins.	Auditory Perception	

Summary of Technical Information

Screening Tests				Achievement Tests									
Harris Graded Word List	Slosson Oral	Dolch Basic Sight Word Test	SRA Reading and Arithmetic Indexes	ABE Student Survey	ABLE*	Fundamental Achievement Series*	TABE*	WRAT	Basic Reading Inventory*	IRPI*	ITED*	MAT	
													Validity:
			X	X	X	X	X	X		X	X	X	1. Is information on validity labeled and readily available?
			X	X		X				X			2. Does the test exhibit validity data for adults?
			X	X	X	X	X	X			X		3. Is the type of validity clearly stated?
													Norms:
				X	X	X	X	X			X	X	1. Is information on standardization available?
				X	X	X	X				X	X	2. Did the sample contain persons similar in background to adult students in N.J.?
				X		X							3. Did the sample contain adults?
													Reliability:
	X		X	X	X	X	X	X	X	X	X	X	1. Is information on realiability available?
	X		X	X	X	X	X	X		X	X	X	2. Is the type of coefficient computed reported?
						X		X			X	X	3. Is a SE of measurement reported?
													Subtests/Parts/Sections:
			X	X	X	X	X	X	X	X	X	X	1. Does the test contain subscores?
			X		X	X	X	X			X	X	2. Are these scores reliable?

*Some information obtained from other than teacher's manual.

Summary of Technical Information

ESL Placement Test 400-600*	ESL Placement Test 100-300*	Examination in Structure	Ilyin Oral Interview*	SWCCEL	Test of Aural Comprehension	Tests of Aural Perception	
X	X		**	X	X	X	Validity: 1. Is information on validity labeled and readily available?
X	X			X	X		2. Does the test exhibit validity data for adults?
X	X			X	X		3. Is the type of validity clearly stated?
X	X		X		X	X	Norms: 1. Is information on standardization available?
X	X		X				2. Did the sample contain persons similar in background to adult ABE/ESL students in New Jersey?
X	X		X		X	X	3. Did the sample contain adults?
X	X		**		X	X	Reliability: 1. Is information on reliability available?
X	X				X	X	2. Is the type of coefficient computed reported?
X					X	X	3. Is a SE of measurement reported?
						X	Subtests/Parts/Sections: 1. Does the test contain subscores?
							2. Are these scores reliable?

*Some information obtained from other than teacher's manual.
**Will be available at a future date.

REFERENCES

Anastasi, Anne. *Psychological Testing*. New York: The Macmillan Company, 1969.

Boy, Angelo V. and Pine, Gerald J. *Client-Centered Counseling in the Secondary School*. Boston: Houghton Mifflin Company, 1963.

Buros, Oscar K., ed. *Reading Tests and Reviews*. Highland Park, New Jersey: The Gryphon Press, 1968.

Buros, Oscar K., ed. *The Seventh Mental Measurements Yearbook*, Volumes 1-2. Highland Park, New Jersey: The Gryphon Press, 1972.

Buros, Oscar K., ed. *The Sixth Mental Measurements Yearbook*. Highland Park, New Jersey: The Gryphon Press, 1965.

Dolch, Edward W. *Teaching Primary Reading*. Champaign, Illinois: Garrard Publishing Company, 1950.

Farr, Roger and Anastasiow. *Tests of Reading and Achievement: A Review and Evaluation*. Newark, Delaware: International Reading Association, 1969.

Farwell, Gail F. and Peters, Herman J., ed. *Guidance Readings for Counselors*. Chicago: Rand McNally Company, 1960.

Fausto, Vincent. *Teaching English as a Second Language to Adults*. New Brunswick, New Jersey: Rutgers Vocational-Technical Curriculum Laboratory in conjunction with the State Department of Education, Division of Vocational Education, October, 1971.

Harris, David P. *Testing English as a Second Language*. New York: McGraw Hill Book Company, 1969.

Helmstadter, G.C. *Principles of Psychological Measurement*. New York: Appleton-Century-Crofts, 1964.

Johnson, Dale D. "The Dolch List Re-examined," *The Reading Teacher*. January, 1971, pp. 449-457.

Joint Committee. *Technical Recommendations for Psychological Tests and Diagnostic Techniques*. Washington, D.C.: American Psychological Association, Inc., 1954.

Katz, Martin, ed. *Selecting an Achievement Test: Principles and Procedures*. Princeton, New Jersey: Educational Testing Service, 1961.

Lado, Robert. *Language Testing: The Construction and Use of Foreign Language Tests*. New York: McGraw-Hill Book Company, 1961.

Mangano, Joseph. "Testing and Test Results." Unpublished Advisory Bulletin, Bureau of Basic Continuing Education, State Department of Education, New York.

McDaniel, H.B.; Eallas, E.; Saum, James A.; and Gilmore, James. Readings in Guidance. New York: Henry Holt and Company, 1969.

National Association of Public Continuing Adult Education. Adult Basic Education: A Guide for Teachers and Teacher Trainers. Washington, D.C.: National Association of Public Continuing Adult Education, 1966.

Thorndike, Robert L. and Hagen, Elizabeth. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc., 1961.

Valette, Rebecca M. Modern Language Testing: A Handbook. New York: Harcourt, Brace, and World, Inc., 1967.